

Error Bounds for Parametric Polynomial Systems with Applications to Higher-Order Stability Analysis and Convergence Rates

G. LI*, B. S. MORDUKHOVICH†, T. T. A. NGHIA‡, and T. S. PHẠM§

September 15, 2015

Dedicated to Terry Rockafellar in honor of his 80th birthday

Abstract

The paper addresses parametric inequality systems described by polynomial functions in finite dimensions, where state-dependent infinite parameter sets are given by finitely many polynomial inequalities and equalities. Such systems can be viewed, in particular, as solution sets to problems of generalized semi-infinite programming with polynomial data. Exploiting the imposed polynomial structure together with powerful tools of variational analysis and semialgebraic geometry, we establish a far-going extension of the Lojasiewicz gradient inequality to the general nonsmooth class of supremum marginal functions as well as higher-order (Hölder type) local error bounds results with explicitly calculated exponents. The obtained results are applied to higher-order quantitative stability analysis for various classes of optimization problems including generalized semi-infinite programming with polynomial data, optimization of real polynomials under polynomial matrix inequality constraints, and polynomial second-order cone programming. Other applications provide explicit convergence rate estimates for the cyclic projection algorithm to find common points of convex sets described by matrix polynomial inequalities and for the asymptotic convergence of trajectories of subgradient dynamical systems in semialgebraic settings.

Keywords: Polynomial Optimization, Generalized Semi-Infinite Programming, Error Bounds, Variational Analysis, Generalized Differentiation, Semialgebraic Functions and Sets, Lojasiewicz Inequality, Second-Order Cone Programming, Higher-Order Stability Analysis, Convergence Rate of Algorithms

AMS Subject Classification: 90C26, 90C31, 90C34, 49J52, 49J53, 26C05

1 Introduction

This paper is largely devoted to *polynomial semi-infinite optimization* and related topics being revolved around deriving *explicit error bounds* for infinite parametric inequality systems with (real) *polynomial data* as well as their various applications to stability analysis in optimization and convergence of algorithms. The imposed polynomial structure allows us to widely use powerful tools of semialgebraic geometry, while

*Department of Applied Mathematics, University of New South Wales, Sydney 2052, Australia. E-mail: g.li@unsw.edu.au. Research of this author was partly supported by the Australian Research Council.

†Department of Mathematics, Wayne State University, Detroit, MI 48201, USA. Email: boris@math.wayne.edu. Research of this author was partly supported by the USA National Science Foundation under grants DMS-1007132 and DMS-1512846 and by the USA Air Force Office of Scientific Research under grant No. 15RT0462.

‡Department of Mathematics and Statistics, Oakland University, Rochester, MI 48306, USA. E-mail: nttran@oakland.edu.

§Department of Mathematics, University of Dalat, Dalat, Vietnam. E-mail: sonpt@dlu.edu.vn. Research of this author was partly supported by the Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant No. 101.04-2013.07.

parametric inequalities naturally call for applying constructions and results of variational analysis and generalized differentiation. Needless to say that the seminal contributions by Terry Rockafellar to variational analysis and optimization are difficult to overstate, and it is our honor to dedicate this paper to him.

The primary attention of this paper is paid to the *parametric inequality systems*

$$S := \{x \in \mathbb{R}^n \mid f_l(x, y) \leq 0 \text{ for all } y \in Y(x), l = 1, \dots, L\}, \quad (1.1)$$

where each function $f_l: \mathbb{R}^n \times \mathbb{R}^m$ as $l \in L$ for the given natural number $L \in \mathbb{N}$ is a *polynomial*, and where $Y: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is a set-valued mapping that is also described by finitely many polynomials via inequality and equality constraints. Systems of type (1.1) naturally arise as feasible solution sets in problems of generalized semi-infinite programming, second-order cone programming, robust optimization, and matrix inequalities with polynomial data; see below for more details and applications.

One of the most important issues associated with the inequality systems (1.1) is establishing the so-called *error bounds*. Given $\bar{x} \in \mathbb{R}^n$, recall that a (local) *error bound* of S with a *Hölder exponent* $\tau \in (0, 1]$ holds for F at \bar{x} if there exist constants $c > 0$ and $\varepsilon > 0$ such that

$$\text{dist}(x, S) \leq c \left(\left[\sup_{\substack{y \in Y(x), \\ 1 \leq l \leq L}} f_l(x, y) \right]_+ \right)^\tau \text{ for all } x \text{ with } \|x - \bar{x}\| \leq \varepsilon, \quad (1.2)$$

where $\text{dist}(x, S)$ signifies the Euclidean distance between x and S , and where $[\alpha]_+ := \max\{\alpha, 0\}$. The supremum in (1.2) is obviously achieved and it can be replaced by ‘max’ if $Y(x)$ is closed and bounded.

The study of error bounds has attracted a lot of attention of many researchers over the years and has found numerous applications to, in particular, sensitivity analysis for various problems of mathematical programming, termination criteria for numerical algorithms, etc. We refer the reader to [32] for an excellent survey in these directions and to the more recent papers [11, 18, 19, 21, 23, 31, 36] with the bibliographies therein. It is worth noting that the major attention in the aforementioned and many other publications on error bounds has drawn to the case of *linear rate* ($\tau = 1$), where this issue is related to metric regularity and subregularity notions in basic variational analysis. Our main interest in this paper concerns *fractional/root* error bounds $0 < \tau < 1$ in (1.2). For the case of *finite* and *fixed* sets $Y(x) \equiv \Omega$ in (1.2), some results in this direction have been obtained in, e.g., [8, 10, 20, 21, 25, 26, 30] with various applications therein.

It is proved [25] in this finite case of $Y(x) \equiv \Omega$, by using the celebrated Łojasiewicz gradient inequality [24], that (1.2) holds with some *unknown* exponent $\tau \in (0, 1)$ for polynomial systems (1.1). Employing advanced techniques of variational analysis, we have recently derived in this case [22] several error bounds with exponents *explicitly determined* by the dimension of the underlying space and the number/degree of the involved polynomials. The techniques and results developed in [22] allowed us to resolve several open questions raised in the literature, which include establishing explicit Hölder error bounds for nonconvex quadratic systems and higher-order semismoothness of the maximum eigenvalue for symmetric tensors.

The primary goal of this paper is to obtain explicit error bounds of type (1.2) for polynomial inequality systems with *infinite* and *variable* sets $Y(x) \subset \mathbb{R}^m$. Besides undoubted importance of these issues for their own sake, we have been motivated by applications of infinite polynomial systems and error bounds for them to higher-order stability and convergence rates of algorithms in optimization-related areas as well as in asymptotic analysis of dynamical systems, where estimates of type (1.2) with infinite sets $Y(x)$ are crucial.

As the reader can see below, deriving error bounds for the case of infinite and variable sets $Y(x)$ in (1.1) is significantly more involved in comparison with our developments for finite systems in [22]. First we present the following three-dimensional example showing that the error bound (1.2) may fail for any $\tau > 0$ for infinite polynomial inequality systems even in the case of constant sets $Y(x) \equiv \Omega$.

Example 1.1 (failure of Hölder error bounds for infinite polynomial systems). Consider the polynomial system of type (1.1) containing only one inequality given by $f: \mathbb{R} \times \mathbb{R}^2 \rightarrow \mathbb{R}$ in the form

$$f(x, y) \leq 0 \text{ with } y \in \Omega, \text{ where } f(x, y) := xy_1 + y_2, y = (y_1, y_2),$$

where the infinite set Ω is constructed as follows. Take the \mathcal{C}^∞ -smooth function of one variable

$$\phi(x) := \begin{cases} e^{-\frac{1}{x^2}} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0 \end{cases}$$

and define the set $\Omega \subset \mathbb{R}^2$ by the conditions

$$\Omega := \{(y_1, y_2) \mid \exists t \in [-0.5, 0.5] \text{ such that } y_1 = \phi'(t) \text{ and } y_2 = \phi(t) - t\phi'(t)\}, \quad (1.3)$$

where ϕ' stands for the derivative of ϕ . Since ϕ is \mathcal{C}^∞ -smooth, the set Ω in (1.3) is nonempty and compact as the image of a compact interval under a continuous mapping. We claim that

$$\max_{y \in \Omega} f(x, y) = \phi(x) \text{ whenever } x \in [-0.5, 0.5]. \quad (1.4)$$

Indeed, for any $x \in [-0.5, 0.5]$ and $y \in \Omega$ there is $t \in [-0.5, 0.5]$ with $y_1 = \phi'(t)$ and $y_2 = \phi(t) - t\phi'(t)$. Thus

$$f(x, y) = x\phi'(t) + \phi(t) - t\phi'(t) = \phi(t) + (x - t)\phi'(t).$$

Applying the second-order Taylor expansion to the function ϕ tells us that

$$\phi(x) = \phi(t) + \phi'(t)(x - t) + \frac{\phi''(a)}{2}(x - t)^2, \text{ for some } a \in [x, t] \subset [-0.5, 0.5].$$

Note that $\phi''(a) = (4a^{-6} - 6a^{-4})e^{-\frac{1}{a^2}} = 2a^{-6}e^{-\frac{1}{a^2}}(2 - 3a^2) \geq 0$, where the last inequality holds due to the choice of $a \in [-0.5, 0.5]$. Hence we get the relationships

$$f(x, y) = \phi(t) + (x - t)\phi'(t) \leq \phi(x) \text{ for all } x \in [-0.5, 0.5], y \in \Omega,$$

which imply the inequality $\max_{y \in \Omega} f(x, y) \leq \phi(x)$ whenever $x \in [-0.5, 0.5]$. On the other hand, it follows from the above constructions of f and Ω that

$$\max_{y \in \Omega} f(x, y) \geq x\phi'(x) + (\phi(x) - x\phi'(x)) = \phi(x) \text{ for each } x \in [-0.5, 0.5],$$

which therefore justifies the claim in (1.4). Having this in mind, consider the set

$$S := \{x \in \mathbb{R} \mid f(x, y) \leq 0 \text{ for all } y \in \Omega\} \quad (1.5)$$

and observe that $S \cap [-0.5, 0.5] = \{x \mid \phi(x) \leq 0\} = \{0\}$. Let us now check that the local Hölder error bound (1.2) fails for (1.5) at $\bar{x} = 0$ for any exponent $\tau > 0$. To see it, take $x_k := \frac{1}{\sqrt{\ln(k+1)}} \rightarrow 0$ as $k \in \mathbb{N}$ and get for all large k that $d(x_k, S) = d(x_k, S \cap [-0.5, 0.5]) = |x_k|$. This allows us to conclude that

$$\frac{d(x_k, S)}{[\max_{y \in \Omega} f(x_k, y)]_+^{\frac{1}{\tau}}} = \frac{|x_k|}{[\phi(x_k)]_+^{\frac{1}{\tau}}} = \frac{(k+1)^{\frac{1}{\tau}}}{\sqrt{\ln(k+1)}} \rightarrow \infty \text{ as } k \rightarrow \infty$$

whenever $\tau > 0$ is chosen, and thus the error bound (1.2) fails for the system S in (1.5).

In what follows we prove that such a situation *does not emerge* if the sets $Y(x)$ in (1.1) are described by

$$Y(x) = \{y \in \mathbb{R}^m \mid g_i(x, y) \leq 0, i = 1, \dots, r, \text{ and } h_j(x, y) = 0, j = 1, \dots, s\}, \quad (1.6)$$

where g_i and h_j are *polynomials*. It is shown in Section 4 that τ in (1.2) is *explicitly calculated* in terms of degrees of the polynomials and dimensions of the spaces in question. The key of our analysis is a new nonsmooth extension of the *Łojasiewicz inequality* to the class of *supremum marginal functions*

$$\phi(x) := \sup_{y \in Y(x)} f(x, y)$$

described by polynomials f and g_i, h_j in (1.6). This is done in Section 3 by using powerful tools of variational analysis and semialgebraic geometry reviewed in Section 2.

Sections 5 and 6 are devoted to applications. In Section 5 we develop *quantitative higher-order stability analysis* for remarkable classes of polynomial optimization problems: *generalized semi-infinite programming*, *optimization of matrix inequalities*, and *second-order cone programming*. Finally, Section 6 contains explicit estimates of *convergence rates* for the *cyclic projection algorithm* to solve feasibility problems for convex sets described by matrix polynomial inequalities and also for asymptotic analysis of *subgradient dynamical systems* governed by maximum functions with polynomial data.

2 Tools of Variational Analysis and Semialgebraic Geometry

This section briefly discusses some tools of generalized differentiation in variational analysis and of semialgebraic geometry widely used in the paper. Throughout this work we deal with finite-dimensional Euclidean spaces labeled as \mathbb{R}^n and endowed by the inner product $\langle x, y \rangle = x^T y$. The symbol $\mathbb{B}_\varepsilon(x)$ (resp. $\overline{\mathbb{B}}_\varepsilon(x)$) stands for the open (resp. closed) ball with center x and radius $\varepsilon > 0$ while \mathbb{B} (resp. $\overline{\mathbb{B}}$) stands for the open (resp. closed) unit ball centered at the origin in \mathbb{R}^n . Given a set $A \subset \mathbb{R}^n$, its interior (resp. boundary, convex hull, and conic convex hull) is denoted by $\text{int } A$ (resp. $\text{bd } A$, $\text{co } A$, and $\text{cone } A$).

Starting with *variational analysis*, recall first two subdifferential notions needed in what follows. The reader can find more information and references in the books [27, 34].

Given a function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ continuous around x , the *proximal subdifferential* of f at x is

$$\partial_P f(x) := \left\{ v \in \mathbb{R}^n \mid \liminf_{\|h\| \rightarrow 0, h \neq 0} \frac{f(x+h) - f(x) - \langle v, h \rangle}{\|h\|^2} > -\infty \right\}. \quad (2.1)$$

The *limiting subdifferential* of f at x (known also as the general, basic or Mordukhovich subdifferential) is

$$\partial f(x) := \{ v \in \mathbb{R}^n \mid \exists x_k \rightarrow x, v_k \rightarrow v \text{ with } v_k \in \partial_P f(x_k), k \in \mathbb{N} \}. \quad (2.2)$$

We clearly have $\partial_P f(x) \subset \partial f(x)$, where the first set may often be empty (while not so in a dense sense), but the second one is nonempty for any locally Lipschitzian function. Furthermore, the set $\partial_P f(x)$ is always convex but may not be closed, while $\partial f(x)$ is closed but may often be nonconvex. Both subdifferentials (2.1) and (2.2) reduce to the gradient $\{\nabla f(x)\}$ for smooth functions and to the subdifferential of convex analysis for convex ones. A significant advantage of the limiting subdifferential (2.2) is *full calculus* in the general nonconvex setting that is based on variational and extremal principles; see [27, 34] for more details.

The major variational notion used below is the *limiting subdifferential slope* of f at x defined via (2.2) by

$$\mathbf{m}_f(x) := \inf \{ \|v\| \mid v \in \partial f(x) \}, \quad (2.3)$$

where $\inf \{\emptyset\} := \infty$. It reduces to the classical gradient slope $\mathbf{m}_f(x) = \|\nabla f(x)\|$ for smooth functions.

Let us next formulate some continuity notions for set-valued mappings $F: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$; see, e.g., [34]. It is said that F is *outer semicontinuous* (o.s.c.) at \bar{x} if for any sequence (x_k, y_k) converging to (\bar{x}, \bar{y}) with $y_k \in F(x_k)$ we have $\bar{y} \in F(\bar{x})$. Further, F is *inner semicontinuous* (i.s.c.) at \bar{x} if for any sequence $x_k \rightarrow \bar{x}$ and any $\bar{y} \in F(\bar{x})$ there are $y_k \in F(x_k)$ as $k \in \mathbb{N}$ satisfying $y_k \rightarrow \bar{y}$. We say that F is o.s.c. or i.s.c. *around* \bar{x} if it has this property at every x in a neighborhood of \bar{x} .

Finally, we present some notions and facts from (real) *semialgebraic geometry* following [6]. It is said that:

- $A \subset \mathbb{R}^n$ is a *semialgebraic set* if it is a finite union of subsets given by

$$\{x \in \mathbb{R}^n \mid f_i(x) = 0, i = 1, \dots, q, \text{ and } f_i(x) > 0, i = q+1, \dots, p\},$$

where all the functions f_i , $i = 1, \dots, p$, are polynomials of some degrees.

- $f: A \rightarrow B$ is a *semialgebraic mapping* if it maps one semialgebraic set $A \subset \mathbb{R}^n$ to another one $B \subset \mathbb{R}^m$ and its graph $\{(x, y) \in A \times B \mid y = f(x)\}$ is a semialgebraic subset of $\mathbb{R}^n \times \mathbb{R}^m$. We say that $f: A \rightarrow B$ is *locally semialgebraic* around \bar{x} if there exists a neighborhood V of the point $(\bar{x}, f(\bar{x}))$ such that the set $V \cap \{(x, y) \in A \times B \mid y = f(x)\}$ is semialgebraic in $\mathbb{R}^n \times \mathbb{R}^m$.

The class of semialgebraic sets is closed under taking finite intersections, finite unions, and complements; furthermore, a Cartesian product of semialgebraic sets is semialgebraic. A major fact concerning the class of semialgebraic sets is given by the following seminal result of semialgebraic geometry.

Tarski-Seidenberg Theorem. *Images of semialgebraic sets under semialgebraic maps are semialgebraic.*

We also need another fundamental result taken from [1, Theorem 4.2], which provides an exponent estimate in the classical Łojasiewicz gradient inequality for polynomials. For brevity we label it as:

Łojasiewicz Gradient Inequality. *Let f be a polynomial on \mathbb{R}^n with degree $d \in \mathbb{N}$. Suppose that $f(0) = 0$ and $\nabla f(0) = 0$. Then there exist constants $c, \varepsilon > 0$ such that for all $x \in \mathbb{R}^n$ with $\|x\| \leq \varepsilon$ we have*

$$\|\nabla f(x)\| \geq c|f(x)|^{1-\tau}, \quad \text{where } \tau = \mathcal{R}(n, d)^{-1} \quad \text{and} \quad \mathcal{R}(n, d) := \begin{cases} 1 & \text{if } d = 1, \\ d(3d - 3)^{n-1} & \text{if } d \geq 2. \end{cases} \quad (2.4)$$

3 Łojasiewicz Inequality for Supremum Marginal Functions

The main aim of this section is extending the Łojasiewicz gradient inequality (2.4) to the following class of (polynomial) *supremum marginal functions* given by

$$\phi(x) := \sup_{y \in Y(x)} f(x, y), \quad (3.1)$$

where $f: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ is a polynomial of degree at most d , and where the set-valued mapping $Y: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is defined in (1.6) by polynomials $g_i: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ for $i = 1, \dots, r$ and $h_j: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ for $j = 1, \dots, s$ of degree at most d . Functions of type (3.1) are intrinsically nonsmooth, and thus deriving a nonsmooth counterpart of (2.4) for them requires the usage of an appropriate subdifferential of ϕ . The reader will see below that a nonsmooth version of the Łojasiewicz inequality for ϕ in terms of the limiting subdifferential slope $\mathbf{m}_\phi(x)$ from (2.3), which replaces the gradient norm in (2.4), plays a key role in establishing our Hölder-type local error bounds and their subsequent applications in this paper.

To proceed in this direction, we have to calculate the limiting subdifferential of ϕ from (3.1) in terms of its initial data, which is not an easy task by taking into account that the sets $Y(x)$ are infinite and variable. When $Y(x) \equiv \Omega$, (3.1) reduces to the *supremum function* for which the most recent subdifferential results can be found in [28]; see also the references therein. When the supremum in (3.1) is replaced by the infimum, we arrive at the class of (infimum) *marginal functions* well investigated in variational analysis [27]. Needless to say that the supremum and infimum operations are essentially different in unilateral analysis and that (lower) subdifferential properties under supremum are significantly more involved and challenging.

We first show that the functions of type (3.1) enjoy the following important properties.

Proposition 3.1 (supremum marginal functions are locally semialgebraic and upper semicontinuous). *Let the sets $Y(x)$ in (1.6) be nonempty and uniformly bounded around some point $\bar{x} \in \mathbb{R}^n$. Then the function ϕ in (3.1) is locally semialgebraic and locally upper semicontinuous (u.s.c.) around \bar{x} .*

Proof. It follows from the Tarski-Seidenberg Theorem and the assumptions made that ϕ is well defined and semialgebraic around \bar{x} . Let us verify that ϕ is u.s.c. around \bar{x} . Select a number $\varepsilon > 0$ and a compact set $C \subset \mathbb{R}^m$ so that $\emptyset \neq Y(x) \subset C$ for all $x \in \mathbb{B}_\varepsilon(\bar{x})$. We claim that ϕ is locally u.s.c. on $\mathbb{B}_\varepsilon(\bar{x})$. Indeed, assume on the contrary that ϕ is not u.s.c. at some $x \in \mathbb{B}_\varepsilon(\bar{x})$ and then get a sequence $x_k \rightarrow x$ and a number $\delta > 0$ such

that $\phi(x_k) > \phi(x) + \delta$. Using the continuity of f , the closedness of $Y(x)$ for all x , and the inclusion $Y(x_k) \subset C$ allows us to find $y_k \in Y(x_k)$ with $\phi(x_k) = f(x_k, y_k)$. By passing to a subsequence we may suppose without loss of generality that $y_k \rightarrow y$. It is easy to check that $y \in Y(x)$ due to the continuity of g_i as $i = 1, \dots, r$ and h_j as $j = 1, \dots, s$ in (1.6). Since f is continuous, we have $\phi(x_k) = f(x_k, y_k) \rightarrow f(x, y) \leq \phi(x)$, which contradicts the assumption that $\phi(x_k) > \phi(x) + \delta$. Thus ϕ is locally u.s.c. around \bar{x} . \square

Next we study the lower semicontinuity (l.s.c.) and Hölder continuity of the function ϕ (3.1) under rather mild assumptions. Define the *argmaximum set* important for our further analysis by

$$M(x) := \{y \in Y(x) \mid \phi(x) = f(x, y)\} \quad (3.2)$$

and recall that $\text{dist}(A; B) := \inf\{\|a - b\| \mid a \in A, b \in B\}$ for $A, B \subset \mathbb{R}^m$.

Proposition 3.2 (continuity of supremum marginal functions). *Given $\bar{x} \in \mathbb{R}^n$, suppose that the sets $Y(x)$ are nonempty and uniformly bounded around \bar{x} . The following assertions hold:*

- (i) *If $\text{dist}(M(\bar{x}); Y(x)) \rightarrow 0$ as $x \rightarrow \bar{x}$, then ϕ is l.s.c. (and hence continuous) at \bar{x} . Thus the i.s.c. property of Y at \bar{x} ensures that ϕ is continuous at this point.*
- (ii) *If there are numbers $\varepsilon, L > 0$ and $\alpha \in (0, 1]$ such that*

$$\text{dist}(M(x_1); Y(x_2)) \leq L\|x_1 - x_2\|^\alpha \text{ for all } x_1, x_2 \in \mathbb{B}_\varepsilon(\bar{x}), \quad (3.3)$$

then ϕ is Hölder continuous on $\mathbb{B}_\varepsilon(\bar{x})$ with order α .

Proof. Since the sets $Y(x)$ are nonempty and uniformly bounded around \bar{x} , there exist a compact set C and a number $\varepsilon > 0$ such that $Y(x) \subset C$ for all $x \in \mathbb{B}_\varepsilon(\bar{x})$. Let us first verify (i) while arguing by contradiction. Suppose that $\text{dist}(M(\bar{x}); Y(x)) \rightarrow 0$ as $x \rightarrow \bar{x}$ and that ϕ is not l.s.c. at \bar{x} . Then there exist a constant $\nu > 0$ and a sequence $x_k \rightarrow \bar{x}$ such that $\phi(x_k) < \phi(\bar{x}) - \nu$. Since $d(M(\bar{x}); Y(x_k)) \rightarrow 0$ as $k \rightarrow \infty$, we find $y_k \in Y(x_k)$ and $\bar{y}_k \in M(\bar{x})$ with $\|y_k - \bar{y}_k\| \rightarrow 0$. Using the boundedness of $\{\bar{y}_k\}_{k \in \mathbb{N}}$ and passing to a subsequence if necessary allow us to find $\bar{y} \in \mathbb{R}^m$ with $\bar{y}_k \rightarrow \bar{y}$ and thus $y_k \rightarrow \bar{y}$. It is easy to see that $\bar{y} \in Y(\bar{x})$ and that $\phi(\bar{x}) = f(\bar{x}, \bar{y}_k) \rightarrow f(\bar{x}, \bar{y})$, which yield in turn that $\bar{y} \in M(\bar{x})$. Hence we have $\phi(x_k) \geq f(x_k, y_k) \rightarrow f(\bar{x}, \bar{y}) = \phi(\bar{x})$ as $k \rightarrow \infty$ while contradicting the assumption on $\phi(x_k) < \phi(\bar{x}) - \nu$. This verifies that ϕ is l.s.c. at \bar{x} and hence its continuity there by Proposition 3.1. The second statement in (i) follows from the first one and the definitions.

To justify now assertion (ii), suppose that (3.3) holds and pick any $x_1, x_2 \in \mathbb{B}_\varepsilon(\bar{x})$. Since the sets $M(x)$ and $Y(x)$ are compact for all $x \in \mathbb{B}_\varepsilon(\bar{x})$, we deduce from (3.3) that there are $y_1 \in M(x_1)$ and $y_2 \in Y(x_2)$ satisfying $\|y_1 - y_2\| \leq L\|x_1 - x_2\|^\alpha$. Taking into account the Lipschitz continuity of the polynomial f on $\overline{\mathbb{B}_\varepsilon(\bar{x})} \times C$ with some constant $\ell > 0$, it follows that

$$\begin{aligned} \phi(x_1) - \phi(x_2) &\leq f(x_1, y_1) - f(x_2, y_2) \leq \ell(\|x_1 - x_2\| + \|y_1 - y_2\|) \\ &\leq \ell(\|x_1 - x_2\| + L\|x_1 - x_2\|^\alpha) = \ell((2\varepsilon)^{1-\alpha} + L)\|x_1 - x_2\|^\alpha. \end{aligned}$$

Similarly we conclude that $\phi(x_2) - \phi(x_1) \leq \ell((2\varepsilon)^{1-\alpha} + L)\|x_1 - x_2\|^\alpha$, which verifies the Hölder continuity of ϕ on $\mathbb{B}_\varepsilon(\bar{x})$ with order α and thus completes the proof of the proposition. \square

Remark 3.3 (effective conditions for continuity of marginal functions). Proposition 3.2(i) tells us that ϕ in (3.1) is continuous at \bar{x} provided that Y is i.s.c. at this point, which surely holds if Y is *locally Lipschitzian* around \bar{x} in the standard/Hausdorff sense with taking into account that the sets $Y(x)$ from (1.6) are closed, nonempty, and uniformly bounded. Necessary and sufficient conditions for the local Lipschitz property of Y around \bar{x} are given in [27, Corollary 4.39] in terms of the gradients of the constraint functions g_i, h_j at (\bar{x}, \bar{y}) for any $\bar{y} \in Y(\bar{x})$. In this way we get effective conditions ensuring the continuity of ϕ at (in fact around) \bar{x} via the initial data of (1.6). On the other hand, [27, Theorem 3.38(iv)] justifies the *Lipschitz continuity* of ϕ around \bar{x} under the i.s.c. property of the argmaximum mapping (3.2) around (\bar{x}, \bar{y}) with some fixed $\bar{y} \in M(\bar{x})$ and the *Lipschitz-like* (Aubin, pseudo-Lipschitz) property of Y around this pair (\bar{x}, \bar{y}) , which is effectively characterized in [27, Corollary 4.39] via the initial data.

To proceed further, we need the following qualification condition imposed at a reference point relative to some (in fact optimal) subset of the constraint set $Y(x)$ in (3.1).

Definition 3.4 (marginal constrained qualification). *Given $\bar{x} \in \mathbb{R}^n$ and $Y(\bar{x})$ from (1.6), we say that the MARGINAL MANGASARIAN-FROMOVITZ CONSTRAINT QUALIFICATION (MMFCQ) holds at \bar{x} relative to some subset $\Omega \subset Y(\bar{x})$ if there is a vector $\xi \in \mathbb{R}^n$ such that*

$$\left[\sum_{i=1}^r \lambda_i \nabla_x g_i(\bar{x}, y) + \sum_{j=1}^s \kappa_j \nabla_x h_j(\bar{x}, y) \right]^T \xi > 0 \text{ for any } y \in \Omega \quad (3.4)$$

whenever Lagrange multipliers $(\lambda, \kappa) \in \mathbb{R}_+^r \times \mathbb{R}^s \setminus \{0\}$ satisfy the conditions

$$\sum_{i=1}^r \lambda_i \nabla_y g_i(\bar{x}, y) + \sum_{j=1}^s \kappa_j \nabla_y h_j(\bar{x}, y) = 0 \text{ and } \lambda_i g_i(\bar{x}, y) = 0, \quad i = 1, \dots, r. \quad (3.5)$$

Remark 3.5 (relationships of MMFCQ with EMFCQ and MFCQ). The defined MMFCQ for the marginal supremum functions (3.1) is motivated by the *extended Mangasarian-Fromovitz constraint qualification* (EMFCQ) introduced in [17] for *generalized semi-infinite programs* (GSIPs) discussed below in Section 5. Reformulating EMFCQ for (3.1), we see that it involves the cost function f and also requires the validity of (3.4) for a larger set of Lagrange multipliers in comparison with MMFCQ; so MMFCQ is weaker than EMFCQ in general. Observe also that when the functions g_i and h_j do not depend on x , the formulated MMFCQ condition means that there is no $(\lambda, \kappa) \in \mathbb{R}_+^r \times \mathbb{R}^s \setminus \{0\}$ satisfying (3.5), which is equivalent to the conventional *Mangasarian-Fromovitz constraint qualification* (MFCQ) on Ω in the sense that

$$\left. \begin{aligned} & \sum_{i=1}^r \lambda_i \nabla g_i(y) + \sum_{j=1}^s \kappa_j \nabla h_j(y) = 0, \\ & \lambda_i \geq 0, \lambda_i g_i(y) = 0, i = 1, \dots, r, \text{ and } \kappa_j \in \mathbb{R}, j = 1, \dots, s \end{aligned} \right\} \implies \lambda_i = 0, \kappa_j = 0 \quad (3.6)$$

for any $y \in \Omega$. As we see in the subsequent sections, the imposed MMFCQ holds automatically for important classes of polynomial optimization and related problems arising in applications.

It is worth emphasizing that the majority of our applications below requires the validity of MMFCQ for the case when Ω in Definition 3.4 is chosen as the *argmaximum set* (3.2) at the reference point. Consider further the standard *Lagrangian function* for (3.1) with the negative sign of γ taken due to the maximization

$$\mathcal{L}(x, y, \gamma, \lambda, \kappa) := -\gamma f(x, y) + \sum_{i=1}^r \lambda_i g_i(x, y) + \sum_{j=1}^s \kappa_j h_j(x, y). \quad (3.7)$$

Observe that any $y \in M(x)$ is a minimizer of the following nonlinear program:

$$\min -f(x, \cdot) \text{ subject to } g_i(x, \cdot) \leq 0, i = 1, \dots, r, \text{ and } h_j(x, \cdot) = 0, j = 1, \dots, s. \quad (3.8)$$

Applying the classical Lagrange multiplier rule in the Fritz John form to (3.8) tells us that the set $FJ(x, y)$ of multipliers $(\gamma, \lambda, \kappa) \in \mathbb{R}_+ \times \mathbb{R}_+^r \times \mathbb{R}^s$ satisfying the conditions

$$\gamma + \sum_{i=1}^r \lambda_i + \sum_{j=1}^s |\kappa_j| = 1, \nabla_y \mathcal{L}(x, y, \gamma, \lambda, \kappa) = 0, \text{ and } \lambda_i g_i(x, y) = 0, i = 1, \dots, r, \quad (3.9)$$

is always nonempty. Given $\alpha \in [0, \infty]$, consider also the set $FJ_\alpha(x, y)$ of $(\gamma, \lambda, \kappa) \in \mathbb{R}_+ \times \mathbb{R}_+^r \times \mathbb{R}^s$ satisfying

$$\gamma + \sum_{i=1}^r \lambda_i + \sum_{j=1}^s |\kappa_j| \leq \alpha, \nabla_y \mathcal{L}(x, y, \gamma, \lambda, \kappa) = 0, \text{ and } \lambda_i g_i(x, y) = 0, i = 1, \dots, r. \quad (3.10)$$

Using the Lagrangian description, we now show that the MMFCQ property relative to $\Omega = M(\bar{x})$ from (3.2) is *robust* in the general setting of Proposition 3.2(i).

Proposition 3.6 (robustness of MMFCQ). *Given $\bar{x} \in \mathbb{R}^n$, assume that the sets $Y(x)$ are nonempty and uniformly bounded around \bar{x} . If MMFCQ holds at \bar{x} relative to $M(\bar{x})$ and if ϕ is l.s.c. at \bar{x} , then there exists $\delta > 0$ such that MMFCQ is satisfied at any point $x \in \mathbb{B}_\delta(\bar{x})$ relative to $M(x)$.*

Proof. Proposition 3.1 tells us that ϕ is continuous at \bar{x} . Supposing on the contrary that MMFCQ is not robust, i.e., there are $x_k \rightarrow \bar{x}$ as $k \rightarrow \infty$ such that MMFCQ fails at x_k relative to $M(x_k)$. Take any $\xi \in \mathbb{R}^n$ and find $y_k \in M(x_k)$ and $(\lambda_k, \kappa_k) \in \mathbb{R}_+^r \times \mathbb{R}^s \setminus \{0\}$ satisfying $\nabla_y \mathcal{L}(x_k, y_k, 0, \lambda_k, \kappa_k) = 0$, $(\lambda_k)_i g_i(x_k, y_k) = 0$ for $i = 1, \dots, r$, and $\nabla_x \mathcal{L}(x_k, y_k, 0, \lambda_k, \kappa_k)^T \xi \leq 0$. Normalization gives us $\|(\lambda_k, \kappa_k)\| = 1$ for all $k \in \mathbb{N}$. Using now the uniform boundedness of $Y(x)$ around \bar{x} , we select subsequences $y_k \rightarrow \bar{y} \in Y(\bar{x})$ and $(\lambda_k, \kappa_k) \rightarrow (\bar{\lambda}, \bar{\kappa})$ with $\|(\bar{\lambda}, \bar{\kappa})\| = 1$. It follows from the continuity of ϕ at \bar{x} and from $\phi(x_k) = f(x_k, y_k)$ that $\phi(\bar{x}) = f(\bar{x}, \bar{y})$, i.e., $\bar{y} \in M(\bar{x})$. Furthermore, passing to limit as $k \rightarrow \infty$ yields $\bar{\lambda}_i g_i(\bar{x}, \bar{y}) = 0$ for $i = 1, \dots, r$, $\nabla_y \mathcal{L}(\bar{x}, \bar{y}, 0, \bar{\lambda}, \bar{\kappa}) = 0$, and $\nabla_x \mathcal{L}(\bar{x}, \bar{y}, 0, \bar{\lambda}, \bar{\kappa})^T \xi \leq 0$. Since $\xi \in \mathbb{R}^n$ was chosen arbitrarily, this contradicts the assumed MMFCQ at \bar{x} relative to $M(\bar{x})$ and thus completes the proof. \square

The next result important for its own sake plays a crucial role in deriving the extended Lojasiewicz inequality for supremum marginal functions. It explicitly evaluates the limiting subdifferential of for such functions via the initial data and the corresponding Lagrange multipliers being significantly different from the preceding results and techniques of [28] even for the constant mapping Y in (3.1) considered therein.

Theorem 3.7 (limiting subgradients of supremum marginal functions). *Given $\bar{x} \in \mathbb{R}^n$, suppose that the sets $Y(x)$ in (1.6) are nonempty and uniformly bounded around \bar{x} , that MMFCQ holds at \bar{x} relative to $M(\bar{x})$, and that ϕ is l.s.c. around \bar{x} . The following assertions hold:*

(i) *There exists $\varepsilon > 0$ such that for any $x \in \mathbb{B}_\varepsilon(\bar{x})$ and $v \in \partial\phi(x)$ we can find $y^{(q)} \in M(x)$ and $(\gamma^{(q)}, \lambda^{(q)}, \kappa^{(q)}) \in FJ_\infty(x, y^{(q)})$, $q = 1, \dots, n+1$, satisfying the conditions*

$$v = \sum_{q=1}^{n+1} \nabla_x \mathcal{L}(x, y^{(q)}, \gamma^{(q)}, \lambda^{(q)}, \kappa^{(q)}) \quad \text{and} \quad \sum_{q=1}^{n+1} \gamma^{(q)} = 1, \quad (3.11)$$

where \mathcal{L} and FJ_∞ are defined in (3.7) and (3.10), respectively.

(ii) *Given $L > 0$, there are positive numbers ε, α such that for any $x \in \mathbb{B}_\varepsilon(\bar{x})$ and $v \in \partial\phi(x) \cap (L\mathbb{B})$ we can find $y^{(q)} \in M(x)$ and $(\gamma^{(q)}, \lambda^{(q)}, \kappa^{(q)}) \in FJ_\alpha(x, y^{(q)})$, $q = 1, \dots, n+1$, satisfying (3.11).*

Proof. It follows from Proposition 3.6 that there is $\varepsilon > 0$ such that MMFCQ holds at any point $x \in \mathbb{B}_\varepsilon(\bar{x})$ relative to $M(x)$. Moreover, Proposition 3.1 and the l.s.c. assumption on ϕ ensure that this function is continuous on $\mathbb{B}_\varepsilon(\bar{x})$. Let us first evaluate the proximal subdifferential (2.1) and then the limiting subdifferential (2.2) of the supremum marginal function (3.1) at each fixed $x \in \mathbb{B}_\varepsilon(\bar{x})$.

Claim 1: *For any proximal subgradient $v \in \partial_P \phi(x)$ we have*

$$(v, -1) \in \text{cone} \left\{ (\nabla_x \mathcal{L}(x, y, \gamma, \lambda, \kappa), -\gamma) \mid y \in M(x), (\gamma, \lambda, \kappa) \in FJ(x, y) \right\} \subset \mathbb{R}^{n+1}. \quad (3.12)$$

To verify (3.12), deduce from (2.1) that there are constants $\eta > 0$ and $\delta \in (0, \varepsilon - \|x - \bar{x}\|)$ satisfying

$$\phi(u) - \phi(x) \geq \langle v, u - x \rangle - \eta \|u - x\|^2 \quad \text{whenever } u \in \mathbb{B}_\delta(x). \quad (3.13)$$

The assumptions imposed on the mapping Y ensure that $\phi(u) = \max\{f(u, y) \mid y \in Y(u)\}$ for all $u \in \mathbb{R}^n$, which tells us by (3.13) that the pair $(x, \phi(x))$ is a local minimizer of the following GSIP:

$$\min_{(u, z)} [z - \langle v, u - x \rangle + \eta \|u - x\|^2] \quad \text{subject to } f(u, y) - z \leq 0 \quad \text{for all } y \in Y(u). \quad (3.14)$$

Applying to (3.14) the necessary optimality condition from [17, Theorem 1.1], we find $p \in \mathbb{N}$, $y^{(q)} \in M(x)$, $(\gamma^{(q)}, \lambda^{(q)}, \kappa^{(q)}) \in FJ(x, y^{(q)})$, $\alpha \geq 0$, and $\mu^{(q)} \geq 0$ for $q = 1, \dots, p$ so that

$$\alpha + \sum_{q=1}^p \mu^{(q)} = 1 \quad \text{and} \quad \alpha(-v, 1) + \sum_{q=1}^p \mu^{(q)} (\nabla_x \mathcal{L}(x, y^{(q)}, \gamma^{(q)}, \lambda^{(q)}, \kappa^{(q)}), -\gamma^{(q)}) = (0, 0). \quad (3.15)$$

To justify (3.12), it remains to show that $\alpha > 0$ in (3.15). Indeed, assuming the contrary tells us that

$$\sum_{q=1}^p \mu^{(q)} = 1 \quad \text{and} \quad \sum_{q=1}^p \mu^{(q)} (\nabla_x \mathcal{L}(x, y^{(q)}, \gamma^{(q)}, \lambda^{(q)}, \kappa^{(q)}), -\gamma^{(q)}) = (0, 0) \quad (3.16)$$

with $(\gamma^{(q)}, \lambda^{(q)}, \kappa^{(q)}) \in FJ(x, y^{(q)})$ for $q = 1, \dots, p$. Consider the set $I := \{q \in \{1, \dots, p\} \mid \mu^{(q)} \neq 0\}$, which is nonempty by $\sum_{q=1}^p \mu^{(q)} = 1$. It follows from (3.16) that $\mu^{(q)} \gamma^{(q)} = 0$ for $q = 1, \dots, p$. Hence $\gamma^{(q)} = 0$ and $(0, \lambda^{(q)}, \kappa^{(q)}) \in FJ(x, y^{(q)})$ for all $q \in I$. Combining this with (3.16) yields

$$\sum_{q \in I} \mu^{(q)} \nabla_x \mathcal{L}(x, y^{(q)}, 0, \lambda^{(q)}, \kappa^{(q)}) = 0,$$

which contradicts the assumed MMFCQ at x relative to $M(x)$ and thus verifies (3.12) for $v \in \partial_P \phi(x)$.

Claim 2: Any limiting subgradient $v \in \partial \phi(x)$ satisfies (3.12) and thus (3.11).

Fix $v \in \partial \phi(x)$. Since ϕ is continuous on $\mathbb{B}_\varepsilon(\bar{x})$, we find by definition (2.2) sequences $x_k \rightarrow x$ and $v_k \in \partial_P \phi(x_k)$ with $v_k \rightarrow v$. Using (3.12) for v_k and applying the Carathéodory theorem to the conic convex hull in (3.12) give us $y_k^{(q)} \in M(x_k)$, $\mu_k^{(q)} \in \mathbb{R}_+$, and $(\gamma_k^{(q)}, \lambda_k^{(q)}, \kappa_k^{(q)}) \in FJ(x_k, y_k^{(q)})$ for $q = 1, \dots, n+1$ such that

$$v_k = \sum_{q=1}^{n+1} \mu_k^{(q)} \nabla_x \mathcal{L}(x_k, y_k^{(q)}, \gamma_k^{(q)}, \lambda_k^{(q)}, \kappa_k^{(q)}) \quad \text{and} \quad \sum_{q=1}^{n+1} \mu_k^{(q)} \gamma_k^{(q)} = 1. \quad (3.17)$$

Let us show that the sequence of $\mu_k := (\mu_k^{(1)}, \dots, \mu_k^{(n+1)})$ is bounded. Arguing by contradiction, suppose that $\sum_{q=1}^{n+1} \mu_k^{(q)} \rightarrow \infty$ and define $\hat{\mu}_k^{(q)} := \mu_k^{(q)} [\sum_{p=1}^{n+1} \mu_k^{(p)}]^{-1}$ for $q = 1, \dots, n+1$. Since the convergent sequence of v_k is bounded as $v_k \rightarrow v$, we deduce from (3.17) that

$$\sum_{q=1}^{n+1} \hat{\mu}_k^{(q)} \nabla_x \mathcal{L}(x_k, y_k^{(q)}, \gamma_k^{(q)}, \lambda_k^{(q)}, \kappa_k^{(q)}) \rightarrow 0 \quad \text{and} \quad \sum_{q=1}^{n+1} \hat{\mu}_k^{(q)} \gamma_k^{(q)} \rightarrow 0 \quad \text{as } k \rightarrow \infty. \quad (3.18)$$

It follows from the boundedness of $y_k^{(q)}$, $\hat{\mu}_k^{(q)}$, and $(\gamma_k^{(q)}, \lambda_k^{(q)}, \kappa_k^{(q)})$ for $q = 1, \dots, n+1$ that some subsequences of them converge to $y^{(q)}$, $\hat{\mu}^{(q)}$, and $(\gamma^{(q)}, \lambda^{(q)}, \kappa^{(q)})$, respectively. Letting $k \rightarrow \infty$ in (3.18) yields

$$\sum_{q=1}^{n+1} \hat{\mu}^{(q)} \nabla_x \mathcal{L}(x, y^{(q)}, \gamma^{(q)}, \lambda^{(q)}, \kappa^{(q)}) = 0 \quad \text{and} \quad \sum_{q=1}^{n+1} \hat{\mu}^{(q)} \gamma^{(q)} = 0 \quad (3.19)$$

with $(\gamma^{(q)}, \lambda^{(q)}, \kappa^{(q)}) \in FJ(x, y^{(q)})$, $y^{(q)} \in M(x)$, and $\sum_{q=1}^{n+1} \hat{\mu}^{(q)} = 1$. Similarly to Claim 1 we arrive at a contradiction with the validity of MMFCQ at x relative to $M(x)$, which therefore shows that the sequence of $\mu_k = (\mu_k^{(1)}, \dots, \mu_k^{(n+1)}) \in \mathbb{R}_+^{n+1}$ is bounded.

By passing to subsequences again, we get that $\mu_k \rightarrow \mu \in \mathbb{R}_+^{n+1}$, $y_k^{(q)} \rightarrow y^{(q)} \in M(x)$, and $(\gamma_k^{(q)}, \lambda_k^{(q)}, \kappa_k^{(q)}) \rightarrow (\gamma^{(q)}, \lambda^{(q)}, \kappa^{(q)}) \in FJ(x, y^{(q)})$ as $k \rightarrow \infty$ for $q = 1, \dots, n+1$. It follows from (3.17) that

$$v = \sum_{q=1}^{n+1} \mu^{(q)} \nabla_x \mathcal{L}(x, y^{(q)}, \gamma^{(q)}, \lambda^{(q)}, \kappa^{(q)}) \quad \text{and} \quad \sum_{q=1}^{n+1} \mu^{(q)} \gamma^{(q)} = 1, \quad (3.20)$$

which gives us (3.12) for each $v \in \partial \phi(x)$ as stated in Claim 2.

To complete the proof of (i), define $\bar{\gamma}^{(q)} := \mu^{(q)} \gamma^{(q)}$, $\bar{\lambda}^{(q)} := \mu^{(q)} \lambda^{(q)}$, and $\bar{\kappa}^{(q)} := \mu^{(q)} \kappa^{(q)}$ and then observe that the inclusion $(\gamma^{(q)}, \lambda^{(q)}, \kappa^{(q)}) \in FJ(x, y^{(q)})$ yields $(\bar{\gamma}^{(q)}, \bar{\lambda}^{(q)}, \bar{\kappa}^{(q)}) \in FJ_\infty(x, y^{(q)})$ and

$$v = \sum_{q=1}^{n+1} \nabla_x \mathcal{L}(x, y^{(q)}, \bar{\gamma}^{(q)}, \bar{\lambda}^{(q)}, \bar{\kappa}^{(q)}), \quad \sum_{q=1}^{n+1} \bar{\gamma}^{(q)} = 1. \quad (3.21)$$

To justify finally (ii), pick any $v \in \partial\phi(x) \cap (L\mathbb{B})$ with $x \in \mathbb{B}_\varepsilon(\bar{x})$ and find by Claim 2 and the Carathéodory theorem such $y^{(q)} \in M(x)$, $\mu^{(q)} \in \mathbb{R}_+$ and $(\gamma^{(q)}, \lambda^{(q)}, \kappa^{(q)}) \in FJ(x, y^{(q)})$ for $q = 1, \dots, n+1$ that (3.20) holds. Taking into account that all the elements above depend on ε and proceeding as in the proof of Claim 2, we deduce that $\mu^{(q)}$, $q = 1, \dots, n+1$, are uniformly bounded when ε is sufficiently small. This means that there are numbers $\alpha, \varepsilon > 0$ such that $0 \leq \mu^{(q)} \leq \alpha$ for any $\mu^{(q)}$ satisfying (3.20) whenever $x \in \mathbb{B}_\varepsilon(\bar{x})$. Defining $\bar{\gamma}^{(q)} := \mu^{(q)}\gamma^{(q)}$, $\bar{\lambda}^{(q)} := \mu^{(q)}\lambda^{(q)}$, and $\bar{\kappa}^{(q)} := \mu^{(q)}\kappa^{(q)}$ as above and then using $(\gamma^{(q)}, \lambda^{(q)}, \kappa^{(q)}) \in FJ(x, y^{(q)})$, we conclude that $(\bar{\gamma}^{(q)}, \bar{\lambda}^{(q)}, \bar{\kappa}^{(q)}) \in FJ_\alpha(x, y^{(q)})$ and that v satisfies (3.21). This clearly ensures the validity of (3.11) and thus finishes the proof of the theorem. \square

Now we are ready to establish a subdifferential extension of the Łojasiewicz Gradient Inequality to supremum marginal functions with polynomial data in terms of the limiting subdifferential slope (2.3).

Theorem 3.8 (Łojasiewicz subgradient inequality for polynomial systems). *Given $\bar{x} \in \mathbb{R}^n$, suppose that all the assumptions of Theorem 3.7 are satisfied for the marginal supremum function ϕ in (3.1). Then there exist positive constants c and ε such that we have the estimate*

$$\mathbf{m}_\phi(x) \geq c|\phi(x) - \phi(\bar{x})|^{1-\tau} \quad \text{for all } x \in \bar{\mathbb{B}}_\varepsilon(\bar{x}) \quad \text{with } \tau = \mathcal{R}(2n + (m+r+s)(n+1), d+2)^{-1}, \quad (3.22)$$

where the constant \mathcal{R} is taken from (2.4).

Proof. It follows from Proposition 3.1 that ϕ is continuous around \bar{x} . Hence there are $\nu, \varepsilon_1 > 0$ such that $|\phi(x) - \phi(\bar{x})| \leq \nu$ for all $x \in \mathbb{B}_{\varepsilon_1}(\bar{x})$. Define $L := \nu^{1-\tau}$ and $F: \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R}^r \times \mathbb{R}^s \rightarrow \mathbb{R}$ by

$$F(x, y, \gamma, \mu, \kappa) := -\gamma f(x, y) + \sum_{i=1}^r \mu_i^2 g_i(x, y) + \sum_{j=1}^s \kappa_j h_j(x, y) \quad (3.23)$$

and observe that F is a polynomial of degree not greater than $d+2$. Define further another polynomial $P: \mathbb{R}^n \times \mathbb{R}^{m(n+1)} \times \mathbb{R}^n \times \mathbb{R}^{r(n+1)} \times \mathbb{R}^{s(n+1)} \rightarrow \mathbb{R}$ of degree not greater than $d+2$ by

$$P(x, y, \gamma, \mu, \kappa) := \sum_{q=1}^n F(x, y^{(q)}, \gamma^{(q)}, \mu^{(q)}, \kappa^{(q)}) + F(x, y^{(n+1)}, 1 - \sum_{q=1}^n \gamma^{(q)}, \mu^{(n+1)}, \kappa^{(n+1)}) + \phi(\bar{x}), \quad (3.24)$$

where $x \in \mathbb{R}^n$, $y = (y^{(1)}, \dots, y^{(n+1)}) \in \mathbb{R}^{m(n+1)}$ with $y^{(q)} \in \mathbb{R}^m$, $\gamma = (\gamma^{(1)}, \dots, \gamma^{(n)}) \in \mathbb{R}^n$ with $\gamma^{(q)} \in \mathbb{R}$, $\mu = (\mu^{(1)}, \dots, \mu^{(n+1)}) \in \mathbb{R}^{r(n+1)}$ with $\mu^{(q)} \in \mathbb{R}^r$, and $\kappa = (\kappa^{(1)}, \dots, \kappa^{(n+1)}) \in \mathbb{R}^{s(n+1)}$ with $\kappa^{(q)} \in \mathbb{R}^s$. For each $\mu = (\mu_1, \dots, \mu_r) \in \mathbb{R}^r$ denote $\mu^{[2]} := (\mu_1^2, \dots, \mu_r^2) \in \mathbb{R}_+^r$ and consider the set

$$\begin{aligned} K(\varepsilon) &:= \left\{ (x, y, \gamma, \mu, \kappa) \in \mathbb{R}^n \times \mathbb{R}^{m(n+1)} \times \mathbb{R}^n \times \mathbb{R}^{r(n+1)} \times \mathbb{R}^{s(n+1)} \right\} \\ &\quad x \in \bar{\mathbb{B}}_\varepsilon(\bar{x}), \gamma \in \mathbb{R}^n, \sum_{q=1}^n \gamma^{(q)} \leq 1, \gamma^{(1)}, \dots, \gamma^{(n)} \geq 0, \\ &\quad y = (y^{(1)}, \dots, y^{(n+1)}), \mu = (\mu^{(1)}, \dots, \mu^{(n+1)}), \kappa = (\kappa^{(1)}, \dots, \kappa^{(n+1)}), \\ &\quad y^{(q)} \in M(x), q = 1, \dots, n+1, \\ &\quad (\gamma^{(q)}, (\mu^{(q)})^{[2]}, \kappa^{(q)}) \in FJ_\alpha(x, y^{(q)}), q = 1, \dots, n, \\ &\quad \left(1 - \sum_{q=1}^n \gamma^{(q)}, (\mu^{(n+1)})^{[2]}, \kappa^{(n+1)} \right) \in FJ_\alpha(x, y^{(n+1)}) \Big\}, \quad \varepsilon \in (0, \varepsilon_2), \end{aligned} \quad (3.25)$$

where positive the numbers α and $\varepsilon_2 \leq \varepsilon_1$ depend on L and are taken from Theorem 3.7(ii). It can be verified that the set-valued mapping $K(\cdot)$ is o.s.c. at $\varepsilon = 0$. Furthermore, for any $(x, y, \gamma, \mu, \kappa) \in K(0)$ we get from (3.10) and (3.25) the equalities

$$\begin{aligned} P(x, y, \gamma, \mu, \kappa) &= - \sum_{q=1}^n \gamma^{(q)} f(x, y^{(q)}) - \left(\left(1 - \sum_{q=1}^n \gamma^{(q)} \right) f(x, y^{(n+1)}) \right) + \phi(\bar{x}) \\ &= - \sum_{q=1}^n \gamma^{(q)} \phi(\bar{x}) - \left(1 - \sum_{q=1}^n \gamma^{(q)} \right) \phi(\bar{x}) + \phi(\bar{x}) = -\phi(\bar{x}) + \phi(\bar{x}) = 0, \end{aligned}$$

where the second one follows from $y^{(q)} \in M(x) = M(\bar{x})$ as $q = 1, \dots, n+1$. Applying the Łojasiewicz Gradient Inequality (2.4) to the polynomial P in (3.24) gives us positive constants η and c_η such that

$$\|\nabla P(x', y', \gamma', \mu', \kappa')\| \geq c_\eta |P(x', y', \gamma', \mu', \kappa')|^{1-\tau} \quad \text{if } \|(x', y', \gamma', \mu', \kappa') - (x, y, \gamma, \mu, \kappa)\| < \eta \quad (3.26)$$

with $\tau := \mathcal{R}(2n + (m+r+s)(n+1), d+2)$. It is easy to see that $K(0)$ is a compact set, and hence we can find $p \in \mathbb{N}$, $(x^{(\nu)}, y^{(\nu)}, \gamma^{(\nu)}, \mu^{(\nu)}, \kappa^{(\nu)}) \in K(0)$, and $\eta_\nu > 0$ for $\nu = 1, \dots, p$ so that

$$K(0) \subset \bigcup_{\nu=1}^p \mathbb{B}_{\eta_\nu}(x^{(\nu)}, y^{(\nu)}, \gamma^{(\nu)}, \mu^{(\nu)}, \kappa^{(\nu)}).$$

This allows us to find a positive number δ ensuring the inclusion

$$K(0) + \delta\mathbb{B} \subset \bigcup_{\nu=1}^p \mathbb{B}_{\eta_\nu}(x^{(\nu)}, y^{(\nu)}, \gamma^{(\nu)}, \mu^{(\nu)}, \kappa^{(\nu)}).$$

Employing the o.s.c. property of $K(\cdot)$ at $\varepsilon = 0$, we find $\varepsilon_3 \in (0, \varepsilon_2)$ so that

$$K(\varepsilon) \subset K(0) + \delta\mathbb{B} \subset \bigcup_{k=1}^p \mathbb{B}_{\eta_k}(x^{(\nu)}, y^{(\nu)}, \gamma^{(\nu)}, \mu^{(\nu)}, \kappa^{(\nu)})$$

for all $\varepsilon \in [0, \varepsilon_3]$. Denoting $c := \min\{1, c_{\eta_1}, \dots, c_{\eta_p}\}$, deduce from (3.26) that

$$\|\nabla P(x, y, \gamma, \mu, \kappa)\| \geq c |P(x, y, \gamma, \mu, \kappa)|^{1-\tau} \quad \text{if } (x, y, \gamma, \mu, \kappa) \in K(\varepsilon). \quad (3.27)$$

Pick now $\varepsilon \in (0, \varepsilon_3)$ and $x \in \overline{\mathbb{B}_\varepsilon(\bar{x})}$. If $m_\phi(x) \geq L$, then (3.22) holds by the choices of L and c . If $m_\phi(x) < L$, then we take any $v \in \partial\phi(x) \cap (L\mathbb{B}) \neq \emptyset$ and deduce from Theorem 3.7(ii) that there are $y^{(q)} \in M(x)$ and $(\gamma^{(q)}, \lambda^{(q)}, \kappa^{(q)}) \in FJ_\alpha(x, y^{(q)})$, $q = 1, \dots, n+1$, satisfying

$$v = \sum_{q=1}^{n+1} \nabla_x \mathcal{L}(x, y^{(q)}, \gamma^{(q)}, \lambda^{(q)}, \kappa^{(q)}) \quad \text{and} \quad \sum_{q=1}^{n+1} \gamma^{(q)} = 1. \quad (3.28)$$

Define next $\mu_i^{(q)} := \sqrt{\lambda_i^{(q)}}$, $i = 1, \dots, r$, and observe the inclusions

$$(\gamma^{(q)}, [\mu^{(q)}]^2, \kappa^{(q)}) \in FJ_\alpha(x, y^{(q)}), \quad q = 1, \dots, n+1.$$

It follows from (3.28) and the definition of $K(\varepsilon)$ in (3.25) that $(x, y, \gamma, \mu, \kappa) \in K(\varepsilon)$, and thus (3.27) holds for this quintuple. Differentiating (3.24) with taking (3.28) into account implies the equalities

$$\begin{aligned} \nabla_x P(x, y, \gamma, \mu, \kappa) &= \sum_{q=1}^{n+1} \nabla_x \mathcal{L}(x, y^{(q)}, \gamma^{(q)}, \lambda^{(q)}, \kappa^{(q)}) = v, \\ \nabla_y P(x, y, \gamma, \mu, \kappa) &= \left(\nabla_{y^{(1)}} \mathcal{L}(x, y^{(1)}, \gamma^{(1)}, \lambda^{(1)}, \kappa^{(1)}), \dots, \nabla_{y^{(n+1)}} \mathcal{L}(x, y^{(n+1)}, \gamma^{(n+1)}, \lambda^{(n+1)}, \kappa^{(n+1)}) \right) = 0, \\ \nabla_\gamma P(x, y, \gamma, \mu, \kappa) &= (f(x, y^{(1)}) - f(x, y^{(n+1)}), \dots, f(x, y^{(n)}) - f(x, y^{(n+1)})) = 0, \\ \nabla_\mu P(x, y, \gamma, \mu, \kappa) &= \left(2\mu_i^{(q)} g_i(x, y^{(q)}) \right)_{i,q} = 0, \quad \text{and} \quad \nabla_\kappa P(x, y, \gamma, \mu, \kappa) = \left(h_j(x, y^{(q)}) \right)_{j,q} = 0, \end{aligned}$$

which ensure by (3.27), (3.28), and (3.24) the relationships

$$\|v\| = \|\nabla P(x, y, \gamma, \mu, \kappa)\| \geq c |P(x, y, \gamma, \mu, \kappa)|^{1-\tau} = c |\phi(x) - \phi(\bar{x})|^{1-\tau} \quad \text{for any } v \in \partial\phi(x) \cap (L\mathbb{B}).$$

Since $m_\phi(x) < L$, this clearly verifies (3.22) and completes the proof of the theorem. \square

Remark 3.9 (sharper estimate for equality polynomial systems). A close look at the proof of Theorem 3.8 reveals that the exponent τ in (3.22) can sharpen to $\tau = \mathcal{R}(2n + (m+s)(n+1), d+1)^{-1}$ in the case when the polynomial system in (1.6) is given by only the equalities $Y(x) = \{y \in \mathbb{R}^m \mid h_j(x, y) = 0, j = 1, \dots, s\}$. This is due to the fact that the polynomials F in (3.23) and P in (3.24) are of degrees at most $d+1$ instead of $d+2$ as in the general case.

4 Error Bounds for Parametric Polynomial Systems

In this section we derive Hölder error bounds with explicit exponents for general polynomial systems of type (1.1), where the underlying state-dependent sets of parameters $Y(x)$ admit the polynomial description (1.6). We also present important specifications of these estimates in some particular cases. The results obtained are far-going extensions of those established in [22] in the case of finite and constant sets $Y(x) \equiv \Omega$. To proceed, we need the following lemma on error bounds for locally Lipschitz functions taken from [31, Corollary 2].

Lemma 4.1 (sufficient condition for error bounds of Lipschitzian functions). *Let $f: \mathbb{R}^n \rightarrow \mathbb{R}$ be continuous around $\bar{x} \in \text{bd } S_f$, where $S_f := \{x \mid f(x) \leq 0\}$. Assume that there are numbers $c, \varepsilon > 0$ such that $\mathbf{m}_f(x) \geq c |f(x)|^{1-\tau}$ for all $x \in \mathbb{R}^n$ with $\|x - \bar{x}\| \leq \varepsilon$ and $x \notin S_f$. Then we have*

$$\text{dist}(x, S_f) \leq \frac{1}{\tau c} [f(x)]_+^\tau \quad \text{whenever } \|x - \bar{x}\| \leq \frac{\varepsilon}{2}.$$

We are now ready to establish the main result of this section. To avoid some technical complexity on dealing with the l.s.c. property of supremum marginal functions as discussed in Section 3, we assume in what follows that the mapping Y in (1.6) is *inner semicontinuous* around the point in question, which is always the case when $Y(x) \equiv \Omega$ for some fixed set $\Omega \subset \mathbb{R}^m$; see, e.g., (4.5) below.

Theorem 4.2 (Hölder error bounds for parametric polynomial systems). *Let $f_l, g_i, h_j: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ with $l = 1, \dots, L$, $i = 1, \dots, r$, and $j = 1, \dots, s$ be polynomial functions of degrees at most d . Suppose that the mapping Y from (1.6) is i.s.c. with nonempty and uniformly bounded values $Y(x)$ around a given point $\bar{x} \in \mathbb{R}^n$ and that MMFCQ from Definition 3.4 holds at this point relative to the argmaximum set*

$$M(\bar{x}) := \left\{ y \in Y(\bar{x}) \mid \max_{\substack{y \in Y(x) \\ 1 \leq l \leq L}} f_l(\bar{x}, y) = f_l(\bar{x}, y) \text{ for some } l = 1, \dots, L \right\}. \quad (4.1)$$

Then there are numbers $c, \varepsilon > 0$ such that we have the local error bound

$$\text{dist}(x, S) \leq c \left(\left[\max_{\substack{y \in Y(x) \\ 1 \leq l \leq L}} f_l(x, y) \right]_+ \right)^\tau \quad \text{for all } x \in \mathbb{R}^n \text{ with } \|x - \bar{x}\| \leq \varepsilon, \quad (4.2)$$

where S is taken from (1.1), and where the Hölder exponent τ in (4.2) is explicitly calculated by

$$\tau = \begin{cases} \mathcal{R}(2n + (m + r + s)(n + 1), d + 2)^{-1}, & L = 1, \\ \mathcal{R}(2n + (m + r + s + 2)(n + 1), d + L + 1)^{-1}, & L \geq 2, \end{cases} \quad (4.3)$$

with the constant \mathcal{R} defined in (2.4). If we suppose in addition that MMFCQ holds at every $x \in \mathbb{R}^n$ relative to the whole set of parameters $Y(x)$ and that Y are i.s.c. on \mathbb{R}^n , then for any compact set $K \subset \mathbb{R}^n$ there is a number $\bar{c} > 0$ such that we have the following global error bound with the same exponent τ :

$$\text{dist}(x, S) \leq \bar{c} \left(\left[\max_{\substack{y \in Y(x) \\ 1 \leq l \leq L}} f_l(x, y) \right]_+ \right)^\tau \quad \text{whenever } x \in K. \quad (4.4)$$

Proof. Consider first the case of $L = 1$, where $\phi(x) = \sup_{y \in Y(x)} f_1(x, y)$ is continuous around \bar{x} by Propositions 3.1 and 3.2(i). Theorem 3.8 gives us numbers c_0, ε_0 such that for all $x \in \overline{\mathbb{B}}_{\varepsilon_0}(\bar{x})$ we have the estimate

$$\mathbf{m}_\phi(x) \geq c_0 |\phi(x) - \phi(\bar{x})|^{1-\tau} \quad \text{with } \tau = \mathcal{R}(2n + (m + r + s)(n + 1), d + 2)^{-1}.$$

Let us justify the existence of $c, \varepsilon > 0$ for which (4.2) holds with τ from (4.3) as $L = 1$. If $\bar{x} \in \text{bd } S$, i.e., $\phi(\bar{x}) = 0$, we get by Lemma 4.1 that (4.3) is fulfilled with $c = \frac{1}{\tau c_0}$ and $\varepsilon = \varepsilon_0/2$. Since this inequality is automatic for $\bar{x} \in \text{int } S$, it remains to examine the case of $\bar{x} \notin S$. It follows from $\bar{x} \notin S = \{x \mid \phi(x) \leq 0\}$ that $\phi(\bar{x}) > 0$ and so $(\phi(\bar{x}))^\tau > 0$. By the continuity of ϕ and $d(\cdot, S)$ we find $M_1, M_2, \varepsilon > 0$ such that

$$(\phi(x))^\tau \geq M_1 \quad \text{and} \quad \text{dist}(x, S) \leq M_2 \quad \text{for all } x \in \overline{\mathbb{B}}_\varepsilon(\bar{x}),$$

and hence $\text{dist}(x, S) \leq M_2 \leq \frac{M_2}{M_1} \phi(x)^\tau$ for all x from this ball. This verifies (4.2) for $L = 1$.

Next we consider the case of $L \geq 2$ and define the *Lagrange interpolation polynomials* $\gamma_l: \mathbb{R} \rightarrow \mathbb{R}$ by

$$\gamma_l(t) := \prod_{j \neq l, j \in \{1, \dots, L\}} \frac{t - j}{l - j}, \quad l = 1, \dots, L,$$

for which $\gamma_l(l) = 1$ as $l = 1, \dots, L$ and $\gamma_l(j) = 0$ as $j \in \{1, \dots, L\}$ with $j \neq l$. Define the functions

$$\begin{aligned} \tilde{f}(x, y, t) &:= \sum_{l=1}^L \gamma_l(t) f_l(x, y), \quad \tilde{g}_i(x, y, t) := g_i(x, y), \quad i = 1, \dots, r, \\ \tilde{h}_j(x, y, t) &:= h_j(x, y), \quad j = 1, \dots, s, \quad \text{and} \quad \tilde{h}_{s+1}(x, y, t) := \prod_{l=1}^L (t - l) \end{aligned}$$

and observe that \tilde{f} is a polynomial on $\mathbb{R}^n \times \mathbb{R}^{m+1}$ with degree not greater than $d + L - 1$. Form now the modified constraint and argmaximum sets by, respectively,

$$\tilde{Y}(x) := \{(y, t) \in \mathbb{R}^m \times \mathbb{R} \mid \tilde{g}_i(x, y, t) \leq 0, \quad i = 1, \dots, r; \quad \tilde{h}_j(x, y, t) = 0, \quad j = 1, \dots, s+1\},$$

$$\tilde{M}(x) := \{(y, t) \in \tilde{Y}(x) \mid \tilde{f}(x, y, t) = \max_{(y, t) \in \tilde{Y}(x)} \tilde{f}(x, y, t)\}$$

while noting that $\tilde{Y}(x) = Y(x) \times \{1, \dots, L\}$ for $Y(x)$ from (1.6) and that

$$\max_{\substack{y \in Y(x) \\ 1 \leq l \leq L}} f_l(x, y) = \max_{(y, t) \in \tilde{Y}(x)} \tilde{f}(x, y, t), \quad S = \{x \in \mathbb{R}^n \mid \tilde{f}(x, y, t) \leq 0 \text{ for all } (y, t) \in \tilde{Y}(x)\}.$$

It can be directly checked that the mapping \tilde{Y} is i.s.c. around \bar{x} with the nonempty and uniformly bounded values and that the MMFCQ property for the modified system holds at \bar{x} relative to $\tilde{M}(\bar{x})$. Applying the result for the case of $L = 1$ obtained above to the system $\tilde{f}, \tilde{g}_i, \tilde{h}_j$ with (n, m, r, s, d) replaced by $(n, m+1, r, s+1, d+L-1)$, we arrive at (4.2) with τ calculated in (4.3) for $L \geq 2$. The last statement of the theorem (4.4) follows from (4.2) by employing the standard compactness arguments. \square

Remark 4.3 (sharper error bounds for equality polynomial systems). Based on the observation in Remark 3.9 and following the proof of Theorem 4.2, we see that the the exponent estimate τ in (4.2) can be sharpen to $\tau = \mathcal{R}(2n + (m + r + s + 2 \min\{L - 1, 1\})(n + 1), d + L)$ if the polynomial system $Y(x)$ in (1.6) contains only the equality constraints.

In the rest of this section we derive specifications of Theorem 4.2 in some particular classes of parametric polynomial systems and discuss the possibility of sharpening Hölder exponents in certain situations.

Error bounds for infinite polynomial inequality systems

Let us examine the case of (1.1) where the sets $Y(x)$ are *independent* of x being reduced to the constant set

$$\Omega := \{y \in \mathbb{R}^m \mid g_i(y) \leq 0, \quad i = 1, \dots, r; \quad h_j(y) = 0, \quad j = 1, \dots, s\} \quad (4.5)$$

described by polynomials with degrees at most d . In particular, such a case covers important models known as *polynomial matrix inequalities*; see below. Observe also that Ω from (4.5) can be treated as an infinite index set in problems of *semi-infinite programming* (SIP); see, e.g., [13]. However, semialgebraic structures of index sets have not been specifically investigated in the usual framework of SIP.

Following the traditional terminology, we say that Ω is *regular* if MFCQ for (4.5) holds on Ω . Let us present examples of two important classes of regular sets described by polynomials.

- **Regularity of discrete sets.** Considering the discrete set $\Omega := \{1, \dots, p\}$, $p \in \mathbb{N}$, we can write it as

$$\Omega = \{y \in \mathbb{R} \mid h(y) = 0\} \quad \text{with} \quad h(y) := (y - 1)(y - 2) \cdots (y - p).$$

Since we obviously have $\nabla h(y) \neq 0$ for all $y \in \Omega$, this set is regular.

• **Regularity of algebraic manifolds.** It follows directly from the definition that the algebraic manifold

$$\Omega := \{y \in \mathbb{R}^m \mid h_1(y) = 0, \dots, h_s(y) = 0\}, \quad s \leq m,$$

is regular provided that $\text{rank}(\nabla h_1(y), \dots, \nabla h_s(y)) = s$ for all $y \in \Omega$.

As an immediate consequence of Theorem 4.2, we arrive at the following extension of the results in [22] from finite to infinite polynomial inequality systems.

Corollary 4.4 (Hölder error bounds for infinite polynomial systems). *Let $f_l: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$, $l = 1, \dots, L$, be polynomials of degrees at most d , let $\Omega \neq \emptyset$ given in (4.5) be bounded and regular, and let S be defined in (1.1) with $Y(x) \equiv \Omega$. Then for any $\bar{x} \in \mathbb{R}^n$ there exist constants $c, \varepsilon > 0$ such that*

$$\text{dist}(x, S) \leq c \left(\left[\max_{\substack{y \in \Omega \\ 1 \leq l \leq L}} f_l(x, y) \right]_+ \right)^\tau \quad \text{whenever } \|x - \bar{x}\| \leq \varepsilon, \quad (4.6)$$

where $\tau = \mathcal{R}(2n + (m + r + s + 2 \min\{L - 1, 1\})(n + 1), d + L + 1)^{-1}$ with the constant \mathcal{R} taken from (2.4). Furthermore, the error bound (4.6) holds globally with some $c > 0$ on any compact set $K \subset \mathbb{R}^n$.

Remark 4.5 (semialgebraicity of index sets is essential for Hölder error bounds). As follows from Example 1.1, the error bound (4.6) can fail if Ω is not semialgebraic. Indeed, this example shows that the Hölder error bound does not hold whenever $\tau \in (0, 1)$ for f and Ω therein. It remains to check that the set Ω from (1.3) is not semialgebraic. Supposing the contrary implies by the Tarski-Seidenberg Theorem that the function $x \mapsto \max_{y \in \Omega} f(x, y)$ is semialgebraic on $[-0.5, 0.5]$, which is not the case.

Next we present a consequence of Corollary 4.4 to polynomial matrix inequalities. Recall that an $(m \times m)$ polynomial matrix inequality of n variables with degree d is represented in the form $P(x) \preceq 0$, where the notation $M \preceq 0$ means that $-M$ is a positive semidefinite matrix, and where $P: \mathbb{R}^n \rightarrow S^m$ is a mapping such that each (i, j) th element of its image $(P(x))_{ij}$, $1 \leq i, j \leq m$, is a real polynomial of degree at most d . We refer the reader to, e.g., [15, 16] for various applications of such systems.

Corollary 4.6 (error bounds for polynomial matrix inequalities without regularity assumptions). *Fix $d \in \mathbb{N}$ and consider the solution set*

$$S_{PMI} := \{x \mid P(x) \preceq 0\}$$

described by the above polynomial matrix inequality. Then for a compact set $K \subset \mathbb{R}^n$ there is $c > 0$ such that

$$\text{dist}(x, S_{PMI}) \leq c \left(\left[\lambda_{\max}(P(x)) \right]_+ \right)^\tau \quad \text{whenever } x \in K,$$

where $\tau = \mathcal{R}(2n + (m + 1)(n + 1), d + 3)^{-1}$ and λ_{\max} denotes the maximum eigenvalue of a symmetric matrix.

Proof. Define $f(x, y) := \sum_{i,j=1}^m y_i (P(x))_{ij} y_j$, which is a real polynomial of degree $d + 2$ on \mathbb{R}^{n+m} . We have

$$P(x) \preceq 0 \iff \left[f(x, y) \leq 0 \text{ for all } y \in \Omega := \{y \in \mathbb{R}^m \mid \|y\|^2 = 1\} \right]. \quad (4.7)$$

Note that the set $\Omega \neq \emptyset$ in (4.7) is bounded and regular. Applying now Corollary 4.4 to the inequality system on the right-hand side of (4.7) and replacing d by $d + 2$ with taking into account Remark 3.9 ensure that for any compact set K we arrive at the relationships

$$\text{dist}(x, S_{PMI}) \leq c \left(\left[\sup_{\|y\|^2=1} f(x, y) \right]_+ \right)^\tau = c \left(\left[\lambda_{\max}(P(x)) \right]_+ \right)^\tau$$

with some constant $c > 0$. This completes the proof of the corollary. \square

Discussion on sharpness of exponent estimations

Finally in this section, we discuss the possibility to sharpen Hölder exponents in the established error bounds for general infinite polynomial systems in some particular settings. Let us first present an example of a polynomial matrix inequality, where the exponent obtained in Corollary 4.6 is not the sharpest one.

Example 4.7 (sharper error bound for matrix inequalities). Let d be an even number. For any $x = (x_1, \dots, x_n)$ define $A_1(x) := x_1^d$ and then, for any $i = 2, \dots, n$,

$$A_i(x) := \begin{pmatrix} -1 & x_i^d \\ x_i^d & -x_{i-1} \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} + x_{i-1} \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} + x_i^d \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Consider the following $(2n-1) \times (2n-1)$ polynomial matrix inequality system:

$$P(x) := \begin{pmatrix} A_1(x) & 0 & 0 & 0 \\ 0 & A_2(x) & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \dots & A_n(x) \end{pmatrix} \preceq 0$$

and get by direct calculations that $S_{PMI} := \{x \in \mathbb{R}^n \mid P(x) \preceq 0\} = \{0\}$ and

$$\begin{aligned} \lambda_{\max}(P(x)) &= \max \{ \lambda_{\max}(A_i(x)) \mid 1 \leq i \leq n-1 \} \\ &= \max \left\{ \max \left\{ \frac{-(1+x_{i-1})}{2} + \sqrt{\frac{(1-x_{i-1})^2}{4} + x_i^{2d}}, x_1^d \right\} \mid 2 \leq i \leq n \right\} \\ &= \max \left\{ \max \left\{ \frac{x_i^{2d} - x_{i-1}}{\frac{(1+x_{i-1})}{2} + \sqrt{\frac{(1-x_{i-1})^2}{4} + x_i^{2d}}}, x_1^d \right\} \mid 2 \leq i \leq n \right\}. \end{aligned}$$

Observe further that for the parametric family $x(t) := (t^{(2d)^{n-1}}, \dots, t^{2d}, t)$, $t \in (0, 1)$, we have $d(x(t), S_{PMI}) = O(t)$ and $\lambda_{\max}(P(x(t))) = t^{d(2d)^{n-1}}$. Thus the exponent of the error bound for $\{x \in \mathbb{R}^n \mid P(x) \preceq 0\}$ at the origin is at most $\tau = \frac{1}{d(2d)^{n-1}}$. Moreover, there is a constant $c_0 > 0$ such that for all x around $\bar{x} = 0$ we have the relationships

$$\begin{aligned} \lambda_{\max}(P(x)) &= \max \left\{ \max \left\{ \frac{x_i^{2d} - x_{i-1}}{\frac{(1+x_{i-1})}{2} + \sqrt{\frac{(1-x_{i-1})^2}{4} + x_i^{2d}}}, x_1^d \right\} \mid 2 \leq i \leq n \right\} \\ &= \max \left\{ \max \left\{ \left[\frac{x_i^{2d} - x_{i-1}}{\frac{(1+x_{i-1})}{2} + \sqrt{\frac{(1-x_{i-1})^2}{4} + x_i^{2d}}} \right]_+, x_1^d \right\} \mid 2 \leq i \leq n \right\} \\ &\geq \max \left\{ \max \left\{ c_0 [x_i^{2d} - x_{i-1}]_+, x_1^d \right\} \mid 2 \leq i \leq n \right\}, \end{aligned}$$

where the second equality is due to $x_1^d \geq 0$ while the last inequality follows as the denominator in the third equality approaches one as $x \rightarrow 0$. Considering next the finite convex polynomial system

$$x_1^d \leq 0 \text{ and } c_0(x_i^{2d} - x_{i-1}) \leq 0 \text{ for } i = 2, \dots, n,$$

note that its solution set is the same as $S_{PMI} = \{0\}$. It is known [8] that for any finite convex polynomial system, where all the involved polynomials have degree at most d_0 on \mathbb{R}^n , a local Hölder error bound holds with the exponent $\frac{2}{(2d_0-1)^{n+1}}$. Hence there are numbers $c, \delta > 0$ such that

$$d(x, S_{PMI}) \leq c \max \left\{ \max \left\{ c_0 [x_i^{2d} - x_{i-1}]_+, x_1^d \right\} \mid 2 \leq i \leq n \right\}^{\frac{2}{(4d-1)^{n+1}}} \leq c \lambda_{\max}(P(x))^{\frac{2}{(4d-1)^{n+1}}}, \quad \|x\| \leq \delta.$$

Thus the *true exponent* of the Hölder error bound at the origin in this example lies within the interval $[\frac{2}{(4d-1)^{n+1}}, \frac{1}{d(2d)^{n-1}}]$. On the other hand, the exponent value calculated in Corollary 4.6 is

$$\tau = \frac{1}{\mathcal{R}(2n^2 + 4n, d + 3)} = \frac{1}{(d + 3)(3d + 6)^{2n^2 + 4n - 1}},$$

which therefore gives us only a lower estimate of the true Hölder exponent in the local error bound.

We now list two specific structures of polynomial systems, where error bounds can be *sharpened*:

- Let $(x, y) \mapsto f_l(x, y)$ be a *concave polynomial* for all $l = 1, \dots, L$, Ω be a compact convex set, and there exist $x_0 \in \mathbb{R}^n$ with $\max_{\substack{y \in \Omega, \\ 1 \leq l \leq L}} f_l(x_0, y) < 0$. It can be directly verified in this case that $x \mapsto \max_{\substack{y \in \Omega, \\ 1 \leq l \leq L}} f_l(x, y)$ is a continuous concave function. Note that a *Lipschitz local error bound* (i.e., with the best possible exponent $\tau = 1$) holds for a concave function satisfying the imposed strict feasibility condition; see [35, Theorem 3].

- Let $x \mapsto f_l(x, y)$ be a *convex polynomial* for each $y \in \mathbb{R}^m$ and $l = 1, \dots, L$, and let Ω be a convex polytope in \mathbb{R}^m . Since in this case the number of extreme points of Ω is finite and is at most $m + 1$, we get the representation $\Omega = \text{co}\{y^{(1)}, \dots, y^{(m+1)}\}$. Noting that a convex function attains its maximum over a convex polytope at its extreme points, it follows that $\max_{\substack{y \in \Omega, \\ 1 \leq l \leq L}} f(x, y) = \max_{1 \leq i \leq m+1} f_l(x, y^{(i)})$. Thus the corresponding explicit error bounds can be simplified and sharpened in this case by using the results from our preceding paper [22] for finite index sets Ω .

5 Higher-Order Stability Analysis for Optimization Problems

In this section we develop some applications of the error bound results from Section 4 to higher-order stability analysis for three important classes in *parametric polynomial optimization* dealing with *infinite* sets of parameters $Y(x)$ from (1.6). In contrast to the majority of publications in parametric optimization, our main emphasis is on *higher-order stability* of set-valued solution maps with *explicit calculating* the corresponding exponents. Let us start with problems of generalized semi-infinite programming.

5.1 Generalized Semi-Infinite Programs with Polynomial Data

Consider the following *polynomial generalized semi-infinite program*:

$$\begin{aligned} (GSIP) \quad & \text{minimize}_{x \in \mathbb{R}^n} \quad p_0(x), \\ & \text{subject to} \quad f_l(x, y) \leq 0 \quad \text{for all } y \in Y(x), \quad l = 1, \dots, L, \end{aligned}$$

where the sets $Y(x) \subset \mathbb{R}^m$ are taken from (1.6), and where all the functions p_0 , f_l , g_i , and h_j are real polynomials with degrees at most d . Along with problem (GSIP) we consider also its *perturbation*

$$\begin{aligned} (GSIP_u) \quad & \text{minimize}_{x \in \mathbb{R}^n} \quad p_0(x, u), \\ & \text{subject to} \quad f_l(x, y) \leq 0 \quad \text{for all } y \in Y(x), \quad l = 1, \dots, L, \end{aligned}$$

where the cost function $p(x, u)$ with $p(x, 0) = p_0(x)$ depends on the perturbation parameter $u \in \mathbb{R}^q$ being polynomial of degree at most d in x and locally Lipschitzian in u . Note that the case of $p(x, u) = p_0(x) - \langle u, x \rangle$ and $Y(x) \equiv \{0\}$ in (GSIP_u) corresponds to the so-called *tilt perturbation* introduced by Poliquin and Rockafellar in [33], where the focus is on *single-valuedness* and *Lipschitz continuity* of the optimal solution map *around* the point in question, which is known as *tilt stability*. In [12, 29] the reader can find recent second-order characterizations of tilt stability in the aforementioned setting of nonlinear programming with C^2 -smooth data and related issues. The *quantitative higher-order stability* for *polynomial* GSIPs established in the next theorem is significantly different: we justify a certain *Hölder continuity* of the *set-valued* solution map *at* the reference point with calculating the corresponding Hölder exponent.

Theorem 5.1 (quantitative higher-order stability for polynomial GSIPs). *Let the feasible set*

$$G := \{x \in \mathbb{R}^n \mid f_l(x, y) \leq 0 \text{ for all } y \in Y(x), l = 1, \dots, L\} \quad (5.1)$$

be nonempty and compact. Assume that the mapping Y in (5.1) is i.s.c. on \mathbb{R}^n with nonempty and uniformly bounded values and that MMFCQ holds on \mathbb{R}^n relative to $Y(x)$. Denote by $S(u)$ the set of optimal solutions to the perturbed problem $(GSIP_u)$. Then for any fixed $\bar{u} \in \mathbb{R}^q$ there are constants $c, \varepsilon > 0$ such that

$$S(u) \subset S(\bar{u}) + c \|u - \bar{u}\|^\tau \text{ if } \|u - \bar{u}\| \leq \varepsilon \text{ with } \tau = \mathcal{R}(2n + (m + r + s + 2)(n + 1), d + L + 2)^{-1}. \quad (5.2)$$

Proof. Consider the optimal value function in $(GSIP_u)$ given by

$$p_*(u) := \inf \{p(x, u) \mid f_l(x, y) \leq 0 \text{ whenever } y \in Y(x), l = 1, \dots, L\}, \quad u \in \mathbb{R}^q,$$

and observe that the solution map $S(u)$ to $(GSIP_u)$ is represented as

$$S(u) = \{x \in \mathbb{R}^n \mid f_l(x, y) \leq 0 \text{ for all } y \in Y(x), l = 1, \dots, L, \text{ and } p(x, u) - p_*(u) \leq 0\}. \quad (5.3)$$

Denote $f_{l+1}(x, y) := p(x, \bar{u}) - p_*(\bar{u})$ and define

$$\phi(x) := \max \left\{ p(x, \bar{u}) - p_*(\bar{u}), \max_{\substack{y \in Y(x), \\ 1 \leq l \leq L}} f_l(x, y) \right\}. \quad (5.4)$$

Applying the error bound from the second part of Theorem 4.2 to the system in (5.3) and (5.4) with L replaced by $L + 1 \geq 2$, we find $c_0 > 0$ such that

$$\text{dist}(x, S(\bar{u})) \leq c_0 [\phi(x)]_+^\tau \text{ for all } x \in G \text{ with } \tau = \mathcal{R}(2n + (m + r + s + 2)(n + 1), d + L + 2)^{-1}.$$

Since $p(x, \cdot)$ is locally Lipschitzian around \bar{u} and G is compact, we find constants $\ell, \varepsilon > 0$ such that

$$|p(x, u_1) - p(x, u_2)| \leq \ell \|u_1 - u_2\| \text{ for all } x \in G \text{ and } u_1, u_2 \in \mathbb{B}_\varepsilon(\bar{u}). \quad (5.5)$$

Denote $c := (2\beta^{-1}\ell)^\tau$ with $\beta := c_0^{-\frac{1}{\tau}} > 0$ and for any $w \in S(u)$ select $z \in S(\bar{u})$ with $\|w - z\| = \text{dist}(w, S(\bar{u}))$. To justify the Hölder stability (5.2), it remains to verify the estimate

$$\|w - z\| \leq c \|u - \bar{u}\|^\tau. \quad (5.6)$$

We consider the case of $w \in S(u) \setminus S(\bar{u})$; otherwise, (5.6) is trivial. Then

$$p(w, \bar{u}) - p_*(\bar{u}) = \phi(w) = [\phi(w)]_+ \geq \beta d(w, S(\bar{u}))^{\frac{1}{\tau}} = \beta \|w - z\|^{\frac{1}{\tau}}.$$

Since $z \in S(\bar{u})$, we get $p(z, \bar{u}) = p_*(\bar{u}) \leq p(w, \bar{u})$, and hence

$$\|w - z\|^{\frac{1}{\tau}} \leq \beta^{-1} (p(w, \bar{u}) - p_*(\bar{u})) \leq \beta^{-1} (p(w, \bar{u}) - p(z, \bar{u})). \quad (5.7)$$

Furthermore, it follows from $w \in S(u)$ that $p(w, u) \leq p(z, u)$, and so (5.5) gives us the relationships

$$\begin{aligned} p(w, \bar{u}) - p(z, \bar{u}) &= (p(w, u) - p(z, u)) + (p(z, u) - p(z, \bar{u})) + (p(w, \bar{u}) - p(w, u)) \\ &\leq (p(z, u) - p(z, \bar{u})) + (p(w, \bar{u}) - p(w, u)) \leq 2\ell \|u - \bar{u}\| \text{ as } w, z \in G \end{aligned}$$

implying together with (5.7) that $\|w - z\|^{\frac{1}{\tau}} \leq \beta^{-1} (p(w, \bar{u}) - p(z, \bar{u})) \leq 2\beta^{-1}\ell \|u - \bar{u}\|$. Therefore

$$\text{dist}(w, S(\bar{u})) = \|w - z\| \leq c \|u - \bar{u}\|^\tau,$$

which justifies (5.6) and thus completes the proof of the theorem. \square

Remark 5.2 (sharper exponents in higher-order stability for equality descriptions.) Based on Remark 4.3, observe from the proof of Theorem 5.1 that the Hölder stability exponent in (5.2) can be sharpened to $\tau = \mathcal{R}(2n + (m + r + s + 2)(n + 1), d + L + 1)$ if the sets $Y(x)$ in $(GSIP)$ are described by the polynomial equalities only: $Y(x) = \{y \mid h_j(x, y) = 0, j = 1, \dots, s\}$.

5.2 Optimization of Polynomial Matrix Inequalities

Next we address stability issues in optimization of *polynomial matrix inequalities*:

$$(PMI) \quad \text{minimize } p_0(x) \text{ subject to } P(x) \preceq 0,$$

where $p_0: \mathbb{R}^n \rightarrow \mathbb{R}$ is a real polynomial of degree not exceeded d , and where the mapping $P: \mathbb{R}^n \rightarrow S^m$ is such that each (i, j) th element of its image $(P(x))_{ij}$, $1 \leq i, j \leq m$, is a real polynomial under the same degree assumption. As in Subsection 5.1, consider the perturbed version of (PMI) :

$$(PMI_u) \quad \text{minimize } p(x, u) \text{ subject to } P(x) \preceq 0,$$

where $u \in \mathbb{R}^q$, $p(x, 0) = p_0(x)$, the degree of $p(\cdot, u)$ is at most d , and $p(x, \cdot)$ is locally Lipschitzian.

Theorem 5.3 (higher-order stability of optimal solution sets to polynomial matrix inequalities).

Assuming that the feasible set $G := \{x \mid P(x) \preceq 0\}$ in (PMI_u) is nonempty and compact and then denoting by $S(u)$ the set of optimal solutions to this problem, for each $\bar{u} \in \mathbb{R}^q$ we find constants $c, \varepsilon > 0$ such that

$$S(u) \subset S(\bar{u}) + c \|u - \bar{u}\|^\tau \text{ if } \|u - \bar{u}\| \leq \varepsilon \text{ with } \tau = \mathcal{R}(2n + (m+3)(n+1), d+4)^{-1}.$$

Proof. Consider the function $f(x, y) := \sum_{i,j=1}^m y_i (P(x))_{ij} y_j$, which is a polynomial of degrees at most $d+2$. We clearly have the equivalence

$$P(x) \preceq 0 \iff f(x, y) \leq 0 \text{ for all } \|y\|^2 = 1.$$

Defining now the constant sets $Y(x)$ by

$$Y(x) \equiv \Omega := \{y \in \mathbb{R}^m \mid \|y\|^2 = 1\},$$

observe that the set Ω is regular. Then applying Theorem 5.1 with $L = 1$, $r = 0$, and $s = 1$, replacing d with $d+2$, and taking into account Remark 5.2 justify the claimed Hölder stability. \square

5.3 Second-Order Cone Programs with Polynomial Data

The class of polynomial optimization problems considers below belongs to *second-order cone programming*

$$(SOCP) \quad \text{minimize } p_0(x) \text{ subject to } \|(f_1^{(l)}(x), \dots, f_m^{(l)}(x))\| \leq f_{m+1}^{(l)}(x), \quad l = 1, \dots, L,$$

where $p_0, f_j^{(l)}: \mathbb{R}^n \rightarrow \mathbb{R}$ for $j = 1, \dots, m+1$ and $l = 1, \dots, L$ are real polynomials with degrees at most d . Besides other areas of applications, problems of this type arise in optimization under *uncertainty* via the *robust optimization approach*. As a simple illustration, let us consider the following quadratic optimization problem under uncertainty in constraint data:

$$\text{minimize } x^T A x + a^T x \text{ subject to } x^T B_l x + b_l^T x + \beta_l \leq 0, \quad l = 1, \dots, L, \quad (5.8)$$

where the triples $(B_l, b_l, \beta_l) \in S^n \times \mathbb{R}^n \times \mathbb{R}$, $l = 1, \dots, L$ are not known in advance. As commonly accepted, suppose that the constraint data belong to the *ellipsoidal uncertainty set*

$$\mathcal{U}_l := \left\{ (B_l^{(0)}, b_l^{(0)}, \beta_l^{(0)}) + \sum_{j=1}^s v_j^l (B_l^{(j)}, b_l^{(j)}, \beta_l^{(j)}) \mid \|(v_l^1, \dots, v_l^s)\| \leq 1 \right\}.$$

The robust optimization approach deals with an associated problem of minimizing an adverse effect of uncertainty. In our case it reads as follows:

$$\text{minimize } x^T A x + a^T x \text{ subject to } x^T B_l x + b_l^T x + \beta_l \leq 0 \text{ for all } (B_l, b_l, \beta_l) \in \mathcal{U}_l, \quad l = 1, \dots, L.$$

By direct calculations we have for each $l = 1, \dots, L$ that

$$\begin{aligned}
& \sup_{(B_l, b_l, \beta_l) \in \mathcal{U}_l} \{x^T B_l x + b_l^T x + \beta_l\} \\
&= x^T B_l^{(0)} x + (b_l^{(0)})^T x + \beta_l^{(0)} + \sup_{\|(v_l^1, \dots, v_l^s)\| \leq 1} \sum_{j=1}^s v_l^j (x^T B_l^{(j)} x + (b_l^{(j)})^T x + \beta_l^{(j)}) \\
&= x^T B_l^{(0)} x + (b_l^{(0)})^T x + \beta_l^{(0)} + \|(x^T B_l^{(1)} x + (b_l^{(1)})^T x + \beta_l^{(1)}, \dots, x^T B_l^{(s)} x + (b_l^{(s)})^T x + \beta_l^{(s)})\|.
\end{aligned}$$

This therefore reduces the uncertainly model to the above framework of (SOCP).

Now we consider the following perturbations of (SOCP) given for all $u \in \mathbb{R}^q$ by

$$(SOCP_u) \quad \text{minimize } p(x, u) \quad \text{subject to } \|(f_1^{(l)}(x), \dots, f_m^{(l)}(x))\| \leq f_{m+1}^{(l)}(x), \quad l = 1, \dots, L,$$

where $p(x, 0) = p_0(x)$, $p(\cdot, u)$ is a polynomial of degree at most d , and $p(x, \cdot)$ is locally Lipschitzian on \mathbb{R}^q .

Theorem 5.4 (higher-order stability of optimal solution sets to polynomial second-order cone programs). *Assuming that the feasible set $G = \{x \mid \|(f_1^{(l)}(x), \dots, f_m^{(l)}(x))\| \leq f_{m+1}^{(l)}(x)\}$ to (SOCP_u) is nonempty and compact, for any $u \in \mathbb{R}^q$ denote by $S(u)$ the set of optimal solutions to this problem. Then given $\bar{u} \in \mathbb{R}^q$, there exist constants $c, \varepsilon > 0$ such that*

$$S(u) \subset S(\bar{u}) + c \|u - \bar{u}\|^\tau \quad \text{if } \|u - \bar{u}\| \leq \varepsilon \quad \text{with } \tau = \mathcal{R}(2n + (mL + L + 2)(n + 1), d + L + 1)^{-1}. \quad (5.9)$$

Proof. For each $l = 1, \dots, L$ we define the function

$$f_l(x, y) := \sum_{j=1}^m y_j^{(l)} f_j^{(l)}(x) - f_{m+1}^{(l)}(x),$$

which is a polynomial with real coefficients on \mathbb{R}^{n+m} of degree at most $d + 1$. Note that

$$\|(f_1^{(l)}(x), \dots, f_m^{(l)}(x))\| \leq f_{m+1}^{(l)}(x) \iff \sup \left\{ \left| \sum_{j=1}^m y_j^{(l)} f_j^{(l)}(x) - f_{m+1}^{(l)}(x) \right| \mid \|(y_1^{(l)}, \dots, y_m^{(l)})\| = 1 \right\} \leq 0$$

whenever $l = 1, \dots, L$. Consider the sets $Y(x) \subset \mathbb{R}^{mL}$ given by

$$Y(x) = \Omega := \{y = (y^{(1)}, \dots, y^{(L)}) \mid y^{(l)} = (y_1^{(l)}, \dots, y_m^{(l)}), \|y^{(l)}\|^2 = 1, l = 1, \dots, L\}, \quad x \in \mathbb{R}^n.$$

Observe that these sets are nonempty and uniformly bounded around any $x \in \mathbb{R}^n$ and that the set Ω is regular. Applying now Theorem 5.1 and Remark 5.2 with $r = 0$ and $s = L$ and with replacing m by mL , we arrive at the claimed Hölder stability in (5.9). \square

6 Convergence Rates in Feasibility and Asymptotic Analysis

The concluding section of the paper presents applications of the obtained error bounds to deriving *explicit convergence rate estimates* in two important frameworks involving polynomial functions. The first issue concerns the cyclic projection algorithm to solve the well-known convex feasibility problem, while the second one deals with asymptotic analysis of subgradient dynamical systems.

6.1 Cyclic Projection Method for Convex Feasibility Problems

A common problem arising in diverse areas of mathematics and applications is to find a point in the intersection of closed convex sets C_l , $l = 1, \dots, L$. This problem is often referred to as the *convex feasibility*

problem. One of the most powerful methods for solving the convex feasibility problem is the so-called *cyclic projection algorithm*, which is formulated as follows. Given finitely many closed convex sets C_1, \dots, C_L in \mathbb{R}^n with $\bigcap_{l=1}^L C_l \neq \emptyset$, pick $x_0 \in \mathbb{R}^n$ and denote $\pi_l := \pi_{C_l}$ for $l = 1, \dots, L$, where π_{C_l} stands for the (unique) Euclidean projection to the set C_l . The sequence of *cyclic projections* $(x_k)_{k \in \mathbb{N}}$ is defined by

$$x_1 := \pi_1 x_0, \dots, x_L := \pi_L x_{L-1}, x_{L+1} := \pi_1 x_L, \dots \quad (6.1)$$

In the case of $L = 2$ the cyclic projection method reduces to the well-known *von Neumann alternating projection method*; see, e.g., [3, 4] and the references therein, where the reader can find various practical applications of the cyclic projection method and related algorithms.

Convergence properties of the cyclic projection algorithm have been examined by many researchers. In particular, the seminal paper by Bregman [9] shows that the sequence $(x_k)_{k \in \mathbb{N}}$ generated by cyclic projections always converges to a point in C . Furthermore, it is proved in [14] that the cyclic projection algorithm converges linearly for regular intersections of convex sets. On the other hand, for convex sets with irregular intersections (e.g., when the intersection is a singleton), the cyclic projection algorithm may not exhibit linear convergence even in simple two-dimensional cases as observed by [2, Example 5.3]. This raises the following basic question: Can we estimate the convergence rate of the cyclic projection algorithm for convex sets with possibly *irregular* intersections of certain particular structures? Quite recently [8] this issue has been investigated in the setting where the convex sets C_l are the so-called *basic semialgebraic* given by

$$C_l := \{x \in \mathbb{R}^n \mid g_i^l(x) \leq 0, i = 1, \dots, r_l\}, \quad l = 1, \dots, L, \quad (6.2)$$

with convex polynomials g_i^l of degrees at most d . It is shown in [8] that for $d > 1$ the sequence of cyclic projections $(x_k)_{k \in \mathbb{N}}$ from (6.1) converges at the rate of

$$\frac{1}{k^\rho} \quad \text{with} \quad \rho := \frac{1}{\min \left\{ (2d-1)^n - 1, 2 \binom{n-1}{\lfloor (n-1)/2 \rfloor} d^n - 2 \right\}},$$

where the symbol $[a]$ stands for the integer part of the number a .

In what follows we establish an explicit convergence rate estimate for the more general case of convex sets described by polynomial matrix inequalities. It is said that a set C is *polynomial matrix inequality representable* if there are numbers $m, d \in \mathbb{N}$ and a mapping $P: \mathbb{R}^n \rightarrow S^m$ such that each (i, j) th element of its image $(P(x))_{ij}$, $1 \leq i, j \leq m$ is a real polynomial with degree at most d and we have the representation $C = \{x \in \mathbb{R}^n \mid P(x) \preceq 0\}$. It is obvious that any basic semialgebraic set (6.2) belongs to this class. The following example shows that the converse is not true.

Example 6.1 (polynomial matrix inequality representable sets may not be basic semialgebraic). Given an odd number d , consider the set

$$C := \left\{ (x_1, x_2) \in \mathbb{R}^2 \mid \begin{pmatrix} -x_1^d & 1 \\ 1 & -x_2^d \end{pmatrix} \preceq 0 \right\} = \{(x_1, x_2) \mid x_1 x_2 \geq 1, x_1 \geq 0, x_2 \geq 0\},$$

which is clearly convex and polynomial matrix inequality representable. Let us show that C cannot be written in the basic semialgebraic form (6.2). Assuming the contrary, take the function $f(x_1, x_2) := x_1$ for which $\inf_{x=(x_1, x_2) \in C} f(x) = 0$. It follows from the polynomial extension of the celebrated (quadratic) Frank-Wolfe theorem that any convex polynomial f bounded from below over a basic semialgebraic convex set attains its minimum; see [5]. This gives us $\bar{x} \in C$ such that $f(\bar{x}) = \inf_{x \in C} f(x) = 0$ for the function f and the set C defined above. It surely contradicts the structures of f and C , and thus verifies the claim.

To derive an explicit convergence rate estimate for the cyclic projection algorithm applied to polynomial matrix inequality representable convex sets, we use the error bound result from Theorem 4.2 and the following abstract result on cyclic projections established in [8] under a certain Hölder condition on convex sets.

Lemma 6.2 (conditional cyclic convergence rate). *Let $C_l \subset \mathbb{R}^n$, $l = 1, \dots, L$, be closed convex sets with $C := \bigcap_{l=1}^L C_l \neq \emptyset$, and let $(x_k)_{k \in \mathbb{N}}$ be the sequence of cyclic projections (6.1) with a starting point x_0 . Impose the following Hölder condition of rank $\tau \in (0, 1]$: for any compact set $K \subset \mathbb{R}^n$ there is a positive number c_0 providing the estimate*

$$di(x, C) \leq c_0 \left(\sum_{l=1}^L \text{dist}(x, C_l) \right)^\tau \quad \text{whenever } x \in K. \quad (6.3)$$

Then there exists $x_\infty \in C$ such that the sequence $(x_k)_{k \in \mathbb{N}}$ converges to x_∞ , and for some $M > 0$ we have

$$\|x_k - x_\infty\| \leq M k^{-\frac{1}{2\tau-1-2}}.$$

Now we are now ready to obtain an explicit convergence rate estimate of the cyclic projection algorithm for polynomial matrix inequality representable convex sets.

Theorem 6.3 (explicit convergence rate of the cyclic projection algorithm). *Let C_l be polynomial matrix inequality representable convex sets given in the form*

$$C_l := \{x \in \mathbb{R}^n \mid A^{(l)}(x) \preceq 0\}, \quad l = 1, \dots, L,$$

where every $A^{(l)}: \mathbb{R}^n \rightarrow S^m$ is a matrix mapping such that each $(A^{(l)}(x))_{ij}$, $1 \leq i, j \leq m$, is a real polynomial with degree at most $d \geq 1$. Given $x_0 \in \mathbb{R}^n$, consider the sequence $(x_k)_{k \in \mathbb{N}}$ of cyclic projections (6.1). Then x_k converges to some $x_\infty \in C$ and there is $M > 0$ ensuring the estimate

$$\|x_k - x_\infty\| \leq M \frac{1}{k^\rho}, \quad k \in \mathbb{N},$$

where $\rho := [2\mathcal{R}(2n + (m+3)(n+1), d+L) - 2]^{-1}$ with the constant \mathcal{R} defined in (2.4).

Proof. Let us first show that for any compact set $K \subset \mathbb{R}^n$ there is $c_0 > 0$ such that condition (6.3) holds with some explicitly calculated $\tau \in (0, 1)$. To proceed, fix a compact set K and define

$$\Omega := \{y \in \mathbb{R}^m \mid \|y\|^2 = 1\} \quad \text{and} \quad f_l(x, y) := \sum_{i,j=1}^m y_i \left(A^{(l)}(x) \right)_{ij} y_j, \quad l = 1, \dots, L.$$

Then we readily have the polynomial set representations

$$C_l = \{x \in \mathbb{R}^n \mid f_l(x, y) \leq 0 \text{ for all } y \in \Omega\} \quad \text{and} \quad C = \left\{x \in \mathbb{R}^n \mid \max_{\substack{y \in \Omega \\ 1 \leq l \leq L}} f_l(x, y) \leq 0\right\}. \quad (6.4)$$

Applying now the error bound result of Theorem 4.2 to the set C from (6.4) with $Y(x) \equiv \Omega$, $r = 0$, and $s = 1$ while taking into account Remark 4.3 and the regularity of the set Ω above, we find c_0 such that

$$\text{dist}(x, C) \leq c_0 \left(\left[\max_{\substack{y \in \Omega \\ 1 \leq l \leq L}} f_l(x, y) \right]_+ \right)^\tau \leq c_0 \left(\sum_{l=1}^L \left[\max_{y \in \Omega} f_l(x, y) \right]_+ \right)^\tau \quad \text{for all } x \in K,$$

where $\tau = \mathcal{R}(2n + (m+3)(n+1), d+L)^{-1}$. For each $l = 1, \dots, L$ define the function $h_l(x) := [\max_{y \in \Omega} f_l(x, y)]_+$, $x \in \mathbb{R}^n$, and observe that it is locally Lipschitzian by Proposition 3.1. Due to the compactness of K we have $M_0 := \max_{x \in K} \sum_{l=1}^L d(x, C_l) < \infty$. Since the set $\widehat{K} := K + M_0 \mathbb{B}$ is also compact, we get $c_l > 0$ ensuring the global Lipschitz condition on this set:

$$|h_l(x_1) - h_l(x_2)| \leq c_l \|x_1 - x_2\| \quad \text{for all } x_1, x_2 \in \widehat{K}, \quad l = 1, \dots, L.$$

Now pick an arbitrary vector $x \in K \subset \widehat{K}$ and select $\widehat{x}_l \in C_l$ such that $\text{dist}(x, C_l) = \|x - \widehat{x}_l\|$. Denoting $c := \max\{c_1, \dots, c_L\} > 0$ and observing that $\widehat{x}_l \in \widehat{K}$ and $h_l(\widehat{x}_l) = 0$ give us the relationships

$$\sum_{l=1}^L \text{dist}(x, C_l) = \sum_{l=1}^L \|x - \widehat{x}_l\| \geq \sum_{l=1}^L \frac{1}{c_l} |h_l(x) - h_l(\widehat{x}_l)| \geq \frac{1}{c} \sum_{l=1}^L |h_l(x)| = \frac{1}{c} \sum_{l=1}^L h_l(x),$$

which ensure therefore the estimates

$$\text{dist}(x, C) \leq c_0 \left(\sum_{l=1}^L h_l(x) \right)^\tau \leq c_0 \left(c \sum_{l=1}^L \text{dist}(x, C_l) \right)^\tau.$$

Thus we arrive at (6.3) and complete the proof of the theorem by using Lemma 6.2. \square

6.2 Subgradient Dynamical Systems

The final piece of this paper is devoted to applications of the obtained error bounds to conducting a quantitative asymptotic analysis of the *subgradient dynamical systems* given by

$$(DS) \quad \begin{aligned} 0 &\in \dot{x}(t) + \partial\phi(x(t)) \quad \text{a.e. on } (0, T), \\ \partial\phi(x(t)) &\neq \emptyset \quad \text{for all } t \in [0, T], \end{aligned}$$

where ϕ is a semialgebraic function on \mathbb{R}^n in the form $\phi(x) := \sup\{f(x, y) \mid y \in \Omega\}$. We suppose in what follows that Ω is a regular, compact, and semialgebraic set defined by

$$\Omega := \{y \in \mathbb{R}^m \mid g_i(y) \leq 0 \text{ for } i = 1, \dots, r \text{ for } h_j(y) = 0, j = 1, \dots, s\} \quad (6.5)$$

and that all the functions f , g_i , and h_j are real polynomials with degrees at most d . By a solution of (DS) we understand any absolutely continuous curve $x: [0, T] \rightarrow \mathbb{R}^n$ satisfying the relationships therein. A trajectory $x(t)$ is called *maximal* if there is no possible extensions of its domain compatible with (DS). Our asymptotic analysis of trajectories for (DS) as $t \rightarrow \infty$ is based on Theorem 3.8 extending the Łojasiewicz inequality to maximum functions and on the following result taken from [7].

Lemma 6.4 (convergence to critical points). *Let $x(t)$ be a bounded maximal trajectory of (DS), where the function ϕ is assumed to be bounded from below on \mathbb{R}^n . Then $x(t)$ is defined on $[0, \infty)$ and converges to a critical point $\bar{x} \in \mathbb{R}^n$ of ϕ , i.e., such that $0 \in \partial\phi(\bar{x})$. Take any $\theta \in [0, 1)$ for which the quantity $\frac{|\phi(x) - \phi(\bar{x})|^\theta}{m_f(x)}$ is bounded for any x around \bar{x} under the convention that $\frac{0}{0} = 1$. Then there are numbers $c, t_0 > 0$ such that*

$$\|x(t) - \bar{x}\| \leq c(t+1)^{-\left(\frac{1-\theta}{2\theta-1}\right)} \quad \text{for all } t \geq t_0.$$

We finally arrive at the constructive rate evaluation of asymptotic convergence for maximal trajectories.

Theorem 6.5 (explicit rate estimate of asymptotic convergence). *Let ϕ be a semialgebraic function on \mathbb{R}^n given by $\phi(x) := \max\{f(x, y) \mid y \in \Omega\}$, where Ω is a semialgebraic set in form (6.5) under the assumptions formulated above. Suppose in addition that ϕ is bounded from below on \mathbb{R}^n . Then any bounded maximal trajectory $x(t)$ for (DS) converges as $t \rightarrow \infty$ to a critical point $\bar{x} \in \mathbb{R}^n$ of ϕ . Furthermore, there are numbers $c, t_0 > 0$, which ensure the estimates*

$$\|x(t) - \bar{x}\| \leq c(t+1)^{-\left(\frac{1-\theta}{2\theta-1}\right)} \quad \text{and} \quad |\phi(x(t)) - \phi(\bar{x})| \leq ct^{-\frac{1}{2\theta-1}} \quad \text{for all } t \geq t_0, \quad (6.6)$$

where $\theta = 1 - \mathcal{R}(2n + (m + r + s)(n + 1), d + 2)^{-1}$ with the constant \mathcal{R} defined in (2.4).

Proof. Let \bar{x} be a critical point of ϕ the existence of which follows from Lemma 6.4. Denote $\tau := \mathcal{R}(2n + (m + r + s)(n + 1), d + 2)^{-1}$ and $\theta := 1 - \tau$. Then Theorem 3.8 tells us that there are numbers $c_0, \varepsilon > 0$ such that $\mathbf{m}_\phi(x) \geq c|\phi(x) - \phi(\bar{x})|^\theta$ for all $x \in \overline{\mathbb{B}}_\varepsilon(\bar{x})$, and therefore

$$\frac{|\phi(x) - \phi(\bar{x})|^\theta}{\mathbf{m}_\phi(x)} \leq c_0^{-1} \quad \text{whenever } x \in \overline{\mathbb{B}}_\varepsilon(\bar{x}). \quad (6.7)$$

Thus the first estimate in (6.6) follows directly from Lemma 6.4. Define now $\gamma: [0, \infty) \rightarrow \mathbb{R}$ by

$$\gamma(t) := \phi(x(t)) - \phi(\bar{x}), \quad t \in [0, \infty),$$

and observe that both functions ϕ and γ are Lipschitz continuous and so a.e. differentiable on $(0, \infty)$. This yields $\partial\phi(x(t)) = \{\nabla\phi(t)\}$ for a.e. $t \in (0, \infty)$; see, e.g., [27, Corollary 1.82]. Then we see that

$$\dot{\gamma}(t) = \langle \nabla\phi(x(t)), \dot{x}(t) \rangle = -\|\nabla\phi(x(t))\|^2 = -\mathbf{m}_\phi(x(t))^2 \leq -c_0^2 |\phi(x(t)) - \phi(\bar{x})|^{2\theta} = -c_0^2 \gamma(t)^{2\theta} \leq 0, \quad (6.8)$$

where the second equality follows from $0 \in \dot{x}(t) + \partial\phi(x(t))$ while the first inequality follows from (6.7). This shows, in particular, that the function $\gamma(t)$ is nonincreasing. Since $x(t) \rightarrow \bar{x}$ and $\gamma(t) \rightarrow 0$ as $t \rightarrow \infty$, we get that either $\gamma(t) \equiv 0$, or $\gamma(t) > 0$ for all large $t > 0$. Integrating (6.8) allows us to find $c, t_0 > 0$ such that the second estimate in (6.6) holds for all $t \geq t_0$, and thus we complete the proof of the theorem. \square

References

- [1] D. D’Acunto and K. Kurdyka, Explicit bounds for the Łojasiewicz exponent in the gradient inequality for polynomials, *Ann. Pol. Math.* **87** (2005), 51–61.
- [2] H. H. Bauschke and J. M. Borwein, Dykstra’s alternating projection algorithm for two sets, *J. Approx. Theory* **79** (1994), 418–443.
- [3] H. H. Bauschke and J. M. Borwein, On projection algorithms for solving convex feasibility problems, *SIAM Rev.* **38** (1996), 367–426.
- [4] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, New York, 2011.
- [5] E. G. Belousov and D. Klatte, A Frank-Wolfe type theorem for convex polynomial programs, *Comput. Optim. Appl.* **22** (2002), 37–48.
- [6] J. Bochnak, M. Coste and M-F. Roy, *Real Algebraic Geometry*, Erg. Math. Grenzgeb. **36**, Springer, Berlin, 1998.
- [7] J. Bolte, A. Daniilidis and A. S. Lewis, The Łojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems, *SIAM J. Optim.* **17** (2007), 1205–1223.
- [8] J. M. Borwein, G. Li and L. Yao, Analysis of the convergence rate for the cyclic projection algorithm applied to basic semi-algebraic convex sets, *SIAM J. Optim.* **24** (2014), 498–527.
- [9] L. M. Bregman, A method of successive projections for finding a common point of convex sets, *Soviet Math. Dokl.* **6** (1965), 688–692.
- [10] S. T. Dinh, H. V. Hà and T. S. Phạm, Hölder-type global error bounds for non-degenerate polynomial systems, preprint, <http://arxiv.org/abs/1411.0859>.
- [11] M. Fabian, R. Henrion, A. Y. Kruger and J. V. Outrata, Error bounds: Necessary and sufficient conditions, *Set-Valued Var. Anal.* **18** (2010), 121–149.

- [12] H. Gfrerer and B. S. Mordukhovich, Complete characterizations of tilt stability in non-linear programming under weakest qualification conditions, to appear in *SIAM J. Optim.*, <http://arxiv.org/abs/1503.04548>.
- [13] M. A. Governa and M. A. López (eds.), *Semi-Infinite Programming: Recent Advances*, Springer, Dordrecht, 2001.
- [14] L. G. Gubin, B. T. Polyak and E. V. Raik, The method of projections for finding the common point of convex sets, *USSR Comput. Math. Math. Physics* **7** (1967), 1–24.
- [15] D. Henrion and J. B. Lasserre, Convergent relaxations of polynomial matrix inequalities and static output feedback, *IEEE Trans. Automat. Control* **51** (2006), 192–202.
- [16] Y. W. Huang, D. P. Palomar and S. Z. Zhang, Lorentz-positive maps and quadratic matrix inequalities with applications to robust MISO transmit beamforming, *IEEE Trans. Signal Process.* **61** (2013), 1121–1130.
- [17] H. Th. Jongen, J.-J. Rückmann and O. Stein, Generalized semi-infinite optimization: A first order optimality condition and examples. *Math. Program.*, **83** (1998), 145–158.
- [18] D. Klatte and W. Li, Asymptotic constraint qualifications and global error bounds for convex inequalities, *Math. Program.* **84** (1999), 137–160.
- [19] A. Y. Kruger, Error bounds and Hölder metric subregularity, to appear in *Set-Valued Var. Anal.* DOI 10.1007/s11228-015-0330-y.
- [20] G. Li, Global error bounds for piecewise convex polynomials, *Math. Program.* **137** (2013), 37–64.
- [21] G. Li and B. S. Mordukhovich, Hölder metric subregularity with applications to proximal point method, *SIAM J. Optim.* **22** (2012), 1655–1684.
- [22] G. Li, B. S. Mordukhovich and T. S. Phạm, New fractional error bounds for polynomial systems with applications to Höderian stability in optimization and spectral theory of tensors, *Math. Program.*, DOI 10.1007/s10107-014-0806-9.
- [23] G. Li and K. F. Ng, Error bounds of generalized D-gap functions for nonsmooth and nonmonotone variational inequality problems, *SIAM J. Optim.* **20** (2009), 667–690.
- [24] M. S. Łojasiewicz, Sur la problème de la division, *Studia Math.* **18** (1959), 87–136.
- [25] X. D. Luo and Z. Q. Luo, Extension of Hoffman’s error bound to polynomial systems, *SIAM J. Optim.* **4** (1994), 383–392.
- [26] Z. Q. Luo, J. S. Pang and D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, 1996.
- [27] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation, I: Basic Theory, II: Applications*, Springer, Berlin, 2006.
- [28] B. S. Mordukhovich and T. T. A. Nghia, Subdifferentials of nonconvex supremum functions and their applications to semi-infinite and infinite programs with Lipschitzian data, *SIAM J. Optim.* **23** (2013), 406–431.
- [29] B. S. Mordukhovich and R. T. Rockafellar, Second-order subdifferential calculus with applications to tilt stability in optimization, *SIAM J. Optim.* **22** (2012), 953–986.
- [30] K. F. Ng and X. Y. Zheng, Global error bounds with fractional exponents, *Math. Program.* **88** (2000), 357–370.
- [31] H. V. Ngai and M. Théra, Error bounds for systems of lower semicontinuous functions in Asplund spaces, *Math. Program.* **116** (2009), 397–427.

- [32] J. S. Pang, Error bounds in mathematical programming, *Math. Program.* **79** (1997), 299–332.
- [33] R. A. Poliquin and R. T. Rockafellar, Tilt stability of a local minimim, *SIAM J. Optim.* **8** (1998), 287–299.
- [34] R. T. Rockafellar and R. J-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.
- [35] H. A. Thi, T. D. Pham and H. V. Ngai, Exact penalty and error bounds in DC programming. *J. Global Optim.*, **52** (2012), 509–535.
- [36] Z. L. Wu and J. J. Ye, On error bounds for lower semicontinuous functions, *Math. Program.* **92** (2002), 301–314.